The current state of Transfer Learning in organizations

Exploratory research on the implementation of Transfer Learning across industries

Author: Sjoerd Rienks

First supervisor: Dr. M. Hilbert
Second supervisor: Prof.dr.ir. J.M.W. Visser
Company supervisor: M. Engbersen Msc

Submitted in partial fulfillment of the requirements for the degree of Master of Science
ICT in Business and the Public Sector

Leiden Institute for Advanced Computer Sciences (LIACS)
Leiden University
The Netherlands
April 17th, 2024

Acknowledgements

This thesis would not have been possible without the support and guidance from multiple people, who I would like to express my gratitude to.

Firstly, to my supervisor Dr. M. Hilbert, for his guidance throughout this writing process and for suggesting this compelling thesis subject. As someone new to the concept of Transfer Learning, he provided me with all the tools and information needed to succeed.

Next, to the whole Digital Transformation team of KPMG, who provided me with an internship opportunity where I could fully focus on my research project. Special appreciation goes out to my company supervisor M. Engbersen, whose assistance was invaluable in navigating the organization on my quest to respondents.

I would also like to thank my family, friends, and housemates, who were always ready to discuss various aspects of my thesis, even when the subject matter was unfamiliar to them.

Finally, I extend my gratitude to the interviewees. I would like to thank them for their time and valuable insights. For confidentiality reasons, the names of the interviewees, together with information about their respective organizations have been kept anonymous.

Abstract

The implementation of Artificial Intelligence (AI) in various industries is haltered because of several factors, including data and computational shortages. Transfer Learning (TL) has the ability to mitigate some factors because it enables organizations to reuse machine learning (ML) models for different tasks. This thesis investigates the adoption of TL across different industries, with a specific focus on the automotive, healthcare, and finance industries. Through qualitative research involving 11 interviews with professionals from diverse backgrounds, the study uncovers challenges and opportunities organizations face when implementing TL in their operations. The findings show that every industry is receptive to the implementation of TL. TL is currently implemented in a reactive use-cased based matter. A big trend related to Natural Language Processing (NLP) and foundation models in TL is discovered, which highlights the need for further industry-specific investigations.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem Statement

Artificial Intelligence (AI) has in the last years become one of the most discussed topics in organizational context (ETO, 2023). Data and computing power has become increasingly available and the implementation of AI in all types of industries has become omnipresent. AI can be referred to as the development of computer systems that can perform tasks which would typically require human intelligence. AI consists of a wide range of technologies aimed at enabling mathematical models to mimic human cognitive function such as learning, reasoning, problem-solving, decision-making, and even demonstrating creativity (Rai et al., 2019). These mimicked human cognitive functions can subsequently be applied across numerous domains. For example, applications for natural language processing (NLP) for customer service chatbots (Ray, 2023) and image recognition for healthcare diagnostics (Davenport & Kalakota, 2019) are being developed and integrated in worldwide organizations rapidly.

Despite the enthusiasm organizations show regarding the implementation of AI systems, there are a numerous amount of inhibitors which can significantly toughen this implementation process. Two of the major ones are the substantial need for a vast amount of training data and the considerable time and energy required to train a new model. Gathering and curating extensive datasets can be resource-intensive and enough data may not always be readily available. Additionally, training AI models demands a lot of computing power and time, which could potentially delay project timelines and increase operational costs (Lins et al., 2021). These inhibitors are especially of importance for smaller businesses or those without easy access to expansive datasets and computational resources. Data and computing power has become increasingly available and the implementation of AI in all types of industries has become omnipresent.

Some of these inhibitors could potentially be mitigated by implementing Transfer Learning (TL), a subdomain of Machine Learning (ML). TL leverages the knowledge and insights gained from one AI task or domain and applies them to another. These tasks could be related to each other, but this is not a requirement (Curreri et al., 2021). In TL for example, a model initially trained for image recognition of automobiles, can be retrained for a different task, such as recognition of bikes. By leveraging pre-trained models, organizations can reduce their data requirements. Since these pre-trained models come equipped with knowledge from large, diverse datasets, it allows them to quickly adapt to specific tasks with relatively small datasets (Zhuang et al., 2020). This is especially the case for the use of Neural Networks (NN), since specific layers in a pre-trained NN can be frozen and reused in a different one (Zhuang et al., 2020). This can make the implementation of AI less resource intensive.

The possible impact of AI, and in particular TL, differs per business industry. If we measure the potential value creation by the willingness of investment in a specific area, then research on private investment in AI by focus area has identified a number of attractive investment opportunities. The biggest investments were made in medical and healthcare; data management, processing, and cloud; fintech, cybersecurity and data protection; and retail (Maslej et al., 2023).

According to research from Chui et al. (2022), 50% of all organizations that participated in their survey had adopted AI in at least one business unit or function. Of the companies that embedded AI in their businesses, across all industries, TL was embedded 16% of the time. With High Tech/Telecom (22%) and financial services (17%) being the highest scoring industries. However, TL was in this research classified as a separate AI field. Technologies as Deep Learning, Computer Vision and Generative Adversarial Networks (GAN) all had their own classification. It is important to note that TL is nothing more than a mean of reaching your AI goals. It could be used to create Deep Learning, Computer Vision and GAN applications. Thus, the actual integration of TL in industries could potentially be higher.

There exists little research on the actual adoption of use cases of TL per industry. Scientific articles exist in which applications of TL for specific industries are discussed, but the actual adoption and implementation of these applications across different industries in real-world scenarios remains underexplored. We know that different categories and classifications of TL exist, but knowledge of which are used in practice and why is something we do not have. This highlights a gap in the practical understanding and utilization of TL methodologies in various industrial contexts.

This is furthermore practically relevant because TL contains a wide spectrum of classifications and tools for organizations to use. This makes it so that it can be difficult for organizations to implement a TL strategy that creates the maximum amount of value for them. The best TL strategy a company can take might depend on their goals, the available data, and in some cases the available pre-trained models that are accessible.

1.2 Research Question

Most existing research mainly discusses the theories, categories, and possible uses of TL. However, not enough detailed, real-world research on how TL is actually used, which TL approaches are being applied, and how TL performs in different industries, currently exists. Because of this, we lack practical knowledge, potentially making it difficult to create TL strategies that are suited to meet the specific needs and challenges of each industry, and to fully take advantage of everything TL has to offer. Research on this topic could enable organizations to create better insights in their data and help maximize their value creation. Therefore, the following research question is proposed:

RQ: *How is TL currently being adopted across different industries and what are challenges and opportunities based on the identified TL approaches?*

To assist the main research question, four additional hypotheses have been crafted:

H1. The adoption of TL varies across different industries, with some industries being more receptive than others.

H2. Factors such as data availability, privacy, industry-specific needs, industry-specific regulations, and technological infrastructure play a significant role in the extent of TL implementation across different industries.

H3. Implemented approaches of TL vary across industries based on specific AI needs.

H4. Opportunities arising from TL approaches are industry-specific. TL can enhance efficiency, accuracy, and innovation in various domains. Identifying and capitalizing on these opportunities require domain-specific expertise and tailored strategies.

1.3 Thesis outline

This thesis starts with a theoretical framework that builds a basic understanding of TL using existing scientific research. The first section explains what TL is, it defines its different categorizations, and it shows its different capabilities. In the following section, the theoretical framework gives a view on how TL is used across different business industries with a specific focus on healthcare, automotive, and finance, while drawing from various studies. This theoretical framework then assists in creating the questions for the interviews. Interviews with industry professionals were conducted to gain practical insights in the application of TL. These discussions are then analyzed, aiming to correlate the industry perspectives with the theoretical underpinnings previously established. In the end, all findings will be brought together in a final summarizing section; the discussion and conclusion. Important side notes of the research are discussed in this section and potential further research directions are proposed.

Chapter 2

Theoretical Framework

2.1 What is Transfer Learning?

2.1.1 Transfer Learning as a psychological concept

The concept of TL comes originally from the studies of psychology and pedagogy. In these fields, TL is sometimes also referred to as Learning Transfer. It encompasses the idea that humans can utilize the knowledge they have from one task to learn a second task more easily (Bray, 1928). Humans are so to say able to transfer their knowledge when learning a new task or skill. For instance, learning to play the guitar is easier when you can already play the piano. The piano and the guitar are obviously not the same instrument, but the knowledge of music theory someone gained while playing the piano can be used to gain knowledge of the guitar. By recognizing the commonality between two things, humans are able to construct a link through which previously learned experiences and knowledge can be transferred to new contexts. Something that is learned in one situation can be used in another. This helps us in learning new skills and information more easily. Many different examples of TL in the real world exist, as shown in figure 2.1.

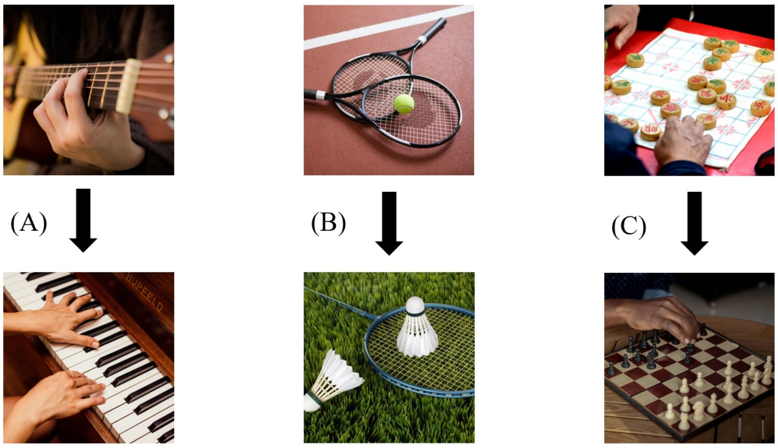


Figure 2.1: Examples of TL in the real world. **(A)** Guitar and piano. **(B)** Tennis and badminton. **(C)** Chinese chess and international chess.

2.1.2 Transfer Learning in Artificial Intelligence

TL is not confined to a psychological concept. When we look at the field of computer science, we see that TL is a specific applicable method of AI. TL in AI leverages the knowledge and insights gained from one task or domain and applies them to another task. These tasks could be related to each other, but this is not a requirement (Curreri et al., 2021). TL can in this way be categorized as a subdomain of ML. In ML, computer systems learn patterns and make predictions or decisions based on data without explicit programming. For traditional ML, one model is trained with data and is focused on only one task. If we want to use a model for a different task, a new model is trained from scratch. In TL however, a model initially trained for a specific task, can be altered so that it is usable for a different one. Figure 2.2 shows the differences between ML and TL.

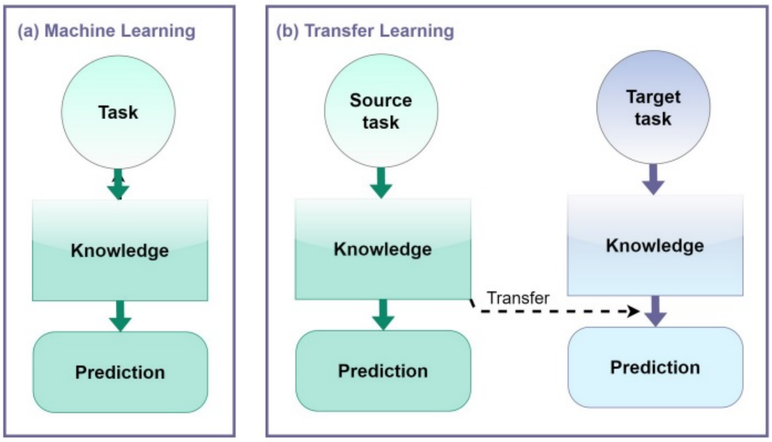


Figure 2.2: Differences between traditional ML **(a)** and TL **(b)** by Curreri et al. (2021).

As an example, we can look at a model that is initially trained for image recognition of automobiles. For TL, it can be adapted so that it is usable for the recognition of bikes. By using pre-trained ML models, we can reap the benefits of AI while reducing our data requirements. Since these pre-trained models come equipped with knowledge from large, diverse datasets, it enables us to quickly adapt to specific tasks with a relatively small dataset (Zhuang et al., 2020). This is especially the case for the use of NN's, since specific layers in a pre-trained NN can be frozen and reused in a different one (Zhuang et al., 2020). More information on how NN's in TL work is given in section 2.5. Overall, TL makes the implementation of AI more accessible to organizations with limited resources.

2.2 Reasons for the use of Transfer Learning

In the domain of AI, TL has become an established technique. Its application in business and research is driven by several reasons which can explain the popularity and effectiveness of TL. Five reasons will be presented in the following section.

2.2.1 Data scarcity

One of the main reasons for the common adoption of TL in AI is the scarcity of data. This scarcity can be classified in two ways. Either the lack of data entirely or the lack of labeled data.

In the current day, organizations and people have more and more data at their disposal from almost every corner of society. Social networks, video surveillance, internet activity, logistics. Enormous amounts of data, so-called 'Big Data' are being generated daily. However, having enough of the right data is often a challenge. For instance, the healthcare industry often deals with data scarcity issues when training ML models for the identification of certain rare diseases (Bansal et al., 2022). Similarly, when data is available, these often lack annotations. In ML, both labeled and unlabeled data can be used for the training of machine learning models, but gathering large enough datasets, labeled or not, can be a time-consuming and costly endeavor (Laurer et al., 2022). Services that can help companies with this exist, take for instance Amazon's SageMaker Ground Truth (Amazon, 2023) or Google's Vertex AI platform (Google, 2023). However, these services provide a costly and time intensive service to label huge amounts of data.

TL mitigates this challenge by allowing models to build upon knowledge gained from pre-existing tasks. Essentially, a pre-trained model learns useful representations from a vast dataset, and this knowledge is then fine-tuned on a smaller, task-specific dataset. This transfer of knowledge significantly reduces the need for extensive labeled data, making AI development more accessible and cost-effective, particularly in industries where collecting abundant labeled data is impractical or prohibitively expensive (Pan & Yang, 2009).

2.2.2 Lack of computational power

The second key driver for the adoption of TL is the issue of a lack computational power. When training deep neural networks from scratch, substantial computing resources are often needed, including high-end CPUs, GPUs and TPUs. These computing resources are however not readily available for the average researcher or company (Verdegem, 2022). Most of the time, only the biggest players in the field can afford to have access to these resources. Companies such as Amazon, Google, Meta, Microsoft and Tesla do have access to big supercomputers and research has shown that companies like that are racing ahead of academia and non-profits in the domain of AI research (Maslej et al., 2023).

TL allows organizations to still reap the benefits of AI while having lesser computational resources by leveraging pre-trained models. This enables the use of smaller and more manageable hardware setups, while still allowing researchers and smaller organizations to create impact (Niu et al., 2020). TL in this case democratizes AI development, making it attainable for a broader audience, including those with limited computational resources.

2.2.3 Maintaining privacy

The third compelling reason behind the adoption of TL in the domain of AI can be its capacity to address privacy concerns effectively. In today's data-driven AI landscape, the loss of privacy is an important concern (Oseni et al., 2021). Moreover, an increasing number of countries around the world are implementing legislation which should guarantee the privacy and data of individuals (Politou et al., 2022). Many industries, including healthcare and finance, are handling sensitive and confidential data that should not be accessible to everyone. TL provides an innovative solution in which AI applications can be build on personal sensitive information, without leaking this data to the public.

One way TL could maintain privacy is by enabling organizations to utilize pre-trained models that have been trained on large, publicly available datasets rather than on personal sensitive data (Y. Wang et al., 2019). These models have already learned valuable features and representations without being exposed to specific user data. This mitigates the risk of data leaks and privacy breaches.

For example, consider a healthcare organization aiming to develop a predictive model for patient outcomes while ensuring the privacy of individual patient records. By employing TL, they can begin with a pre-trained model that has learned from publicly available medical datasets, which does not contain sensitive patient information. They can then fine-tune this model on their own patient data, keeping individual patient records confidential.

In addition, TL is crucial for techniques like federated learning which enable multiple parties to collaboratively train models without sharing raw data. This approach allows organizations to pool their knowledge and resources while preserving data privacy (Xu et al., 2022). It's particularly valuable in scenarios where data cannot be centralized due to privacy regulations or security concerns.

In summary, maintaining privacy is a strong reason for the adoption of TL in AI. It empowers organizations to leverage the power of AI while upholding stringent privacy standards, ensuring that sensitive information remains confidential and protected. This not only aligns with regulatory requirements but also builds trust among users and stakeholders in an increasingly data-conscious world.

2.2.4 Lesser environmental impact

TL is renowned in the field of AI for being an exceptionally efficient and effective technique. One compelling reason contributing to its popularity is its lower environmental impact, which is increasingly crucial in our environmentally conscious society.

In traditional ML, training a model from scratch requires enormous computational resources and energy consumption. Vast amounts of data must be processed and analyzed, which involves extensive calculations performed by powerful hardware over extended periods. This not only leads to high costs but also a significant carbon footprint due to the energy-intensive nature of these computations (Tamburrini, 2022).

To combat this huge environmental impact, Transfer Learning allows the leveraging of pre-trained models that have already undergone extensive training phases. This enables the adaptation of these models to new tasks with less data and computational time. Since the models have acquired knowledge from previous tasks, the process becomes considerably more efficient, reducing the necessary energy consumption and thereby lowering the overall environmental impact (Patterson et al., 2021).

Furthermore, TL promotes model reusability and reduces redundancy in model training, thus saving resources and mitigating unnecessary environmental strains (Patterson et al., 2021). This eco-friendliness, coupled with efficiency, makes Transfer Learning a preferred choice in the sustainable development and application of AI technologies.

An important side note to make however, is the greenhouse emissions made by the organizations developing, training and maintaining the large pre-trained foundation models (Tamburrini, 2022). An organization that utilizes these models might have a lesser environmental impact, but the overall environmental impact of the AI landscape might grow.

2.2.5 Personalized AI

Lastly, TL could be an essential technique in AI for the creation of more personalized AI models. Traditional ML approaches often require large amounts of data to train models that are generalized for a wide audience. However, this one-size-fits-all strategy does not cater to the specific needs or preferences of individual users or small groups. TL can help to enable organizations in developing models that are more personalized.

With TL, a model that has been pre-trained on a large and diverse dataset can be fine-tuned with a smaller and more personalized dataset that reflects the unique characteristics or preferences of a specific user or group. This is for example possible in the development of personalized language models (Yoon et al., 2017). This means that rather than starting the learning process from zero, the model adapts pre-existing knowledge, significantly reducing the need for vast amounts of personalized data.

The ability of TL to accommodate the heterogeneity of individuals is also of great value in a fields such as healthcare, where personalized models can lead to better patient outcomes by considering personal medical history, genetics, and lifestyle factors in diagnostic and treatment plans. For example in the creation of on-device personalized stress modeling NN's, which are developed with TL (Woodward et al., 2020).

By using TL, researchers and organizations have the ability to create AI systems that are not just powerful and predictive, but also deeply customized, offering a level of personalization that was previously difficult or too resource-intensive to achieve. This capability makes it so that TL can be of great influence in making AI more available and applicable to everyone.

2.3 Overview of Transfer Learning Capabilities

TL as a subdomain of ML is not restricted to one application. In fact, TL is applicable to every domain where traditional ML is also applicable. This means that TL is able to be used in a lot of different aspects. In the following section, an overview of the most used TL capabilities and applications is given together with some practical examples.

2.3.1 Computer vision

The first domain in which TL offers great potential is computer vision. Computer vision is a field of computer science aimed at enabling computers to interpret, analyze, and make decisions based on visual data (Voulodimos et al., 2018). The goal is to essentially allow computers to 'see' and understand the content of digital images and videos. It combines the technologies of ML, artificial intelligence, and signal processing to replicate the capabilities of human vision in machines. This allows for the creation of applications such as image and video recognition, object detection, image restoration and image style transfers. This technology is also indispensable for applications such as autonomous vehicles and robotics.

TL is already widely adopted in this domain and is proving to be very capable as a means to developing better AI models. For instance, a model trained on a generic dataset like ImageNet, rich with millions of images across thousands of categories, can be fine-tuned to specialize in specific domains such as medical imaging (H. E. Kim et al., 2022), aerial survey analytics (Liang et al., 2016), or custom object detection (Zhuang et al., 2020). This enables the model to leverage the hierarchical feature representations learned from the vast pre-trainining data, thereby requiring less data and computational resources to adapt to the specific challenges of a new task . Thus, in this situation, transfer learning can act as a catalyst in the development of robust computer vision models, making the possibilities that AI offer more accessible to various industries, domains, and applications.

A practical example of TL in computer vision is research conducted by Xie et al. (2016). The researchers were able to use TL as a technique to map remote poverty data in developing countries. Due to a lack of reliable poverty data in these countries, it is difficult to develop reliable ML models. However, the researchers overcame this hurdle by ingeniously leveraging remote sensory and satellite imagery through TL. In doing so, they were able to extract and analyze meaningful data pertinent to poverty alleviation, paving the way for more informed and strategic socio-economic interventions in the developing countries under study.

2.3.2 Natural Language Processing

The second domain in which TL can be used is the domain of NLP. NLP is a subdomain of AI research in which we try to give computers the ability to understand text and spoken words in the same way as humans do (Khurana et al., 2023). NLP utilizes statistical analysis, ML, and deep learning to enable computers to understand, interpret, and generate human language, producing outputs that are indistinguishable from those created by humans.

The most well-known publicly accessible application for NLP is most likely ChatGPT. ChatGPT is a chatbot made by OpenAI, a company specialized in AI reserach. Users can ask questions, converse or generate text based on written prompts (Leiter et al., 2023). For this, ChatGPT utilizes the Generative Pre-trained Transformer model (GPT). GPT is an enormous pre-trained neural network designed for conversation and text-generation.

TL can play a significant role in enhancing the capabilities of NLP models, which makes it an indispensable tool in the world of NLP. One of the significant benefits of TL in NLP is its ability to leverage pre-trained models, such as GPT, that have already been exposed to enormous amounts of text data. Since these pre-trained models come equipped with an already rich understanding of language syntax, semantics, and linguistic patterns and nuances (Iman et al., 2023). When applied to a new, specific NLP task or application, the models can be fine-tuned to adapt to the particular requirements of the task, enabling them to perform exceptionally well even with limited domain-specific data. This again saves an organization time and resources when developing a NLP model. By not having to start from scratch, applications can be build faster, while still maintaining the same amount of accuracy.

In the context ChatGPT and OpenAI, TL enables us to create more robust and nuanced conversation capabilities. OpenAI already offers a platform in which users can fine-tune the GPT model with their own data so that the model can efficiently handle a specialized array of questions and prompts, thereby producing responses that are coherent, contextually relevant, and human-like (OpenAI, 2023). With this, organizations can for instance create a personalized chatbot that can assist in organization specific tasks.

2.3.3 Audio processing

The third domain where TL can play a role is the processing of many different kinds of audio. Applications in which AI play a role such as speech recognition and general audio classification could definitely benefit from TL techniques.

Voices are next to written text also a part of the expression of the human language. In this domain we can define multiple applications of AI. Automatic Speech Recognition (ASR) is the process of enabling a computer to recognize and understand human generated speech. Another application is the reversal of this process, which is called speech synthesis. In this application we try to make computers generate human like speech. If we do this based on textual input it is called Text To Speech (TTS).

A problem faced in ASR is that every human talks in a different way. The same sentence always sounds a little different when it is spoken by different people. this makes it difficult for machines to understand spoken language. And when we factor in all the different accents people have that speak the same language, availability of data can quickly become a problem. Making a model perform well on a language in all kinds of accents is a challenge. Research has shown that TL can mitigate these kind of challenges (Das et al., 2021).

When we look at the classification of audio data, TL could also be a mitigator of challenges. Research conducted by Choi et al. (2017) describes a TL approach to the classification of music. By using a pre-trained neural network with specific weights, the researchers used the pre-trained weights as a feature extractor for target tasks. In this way, they were able to classify unlabeled music data on different labels such as mood, genre and era. They show that TL can be beneficial to the classification of audio data if the target task does not have enough (labeled) data to successfully train a traditional ML model.

2.3.4 Robotics

The fourth domain in which TL can assist is robotics. Robotics is an interdisciplinary branch of engineering and science which spans over multiple fields such as mechanical engineering, electrical engineering, computer science, and others. Robotics deals with the design, construction, operation, and use of robots, as well as the development of software for their control, sensory feedback, and information processing. These technologies are used to develop machines that can substitute for humans and replicate human actions.

AI can act as the brain of modern robotics, giving robots to ability to make autonomous decisions. Through AI, robots acquire the ability to perceive their environment, process and analyze sensory input, and execute complex tasks with minimal to no human intervention. AI integrates ML and reasoning, allowing robots to learn from experiences and adapt to various situations. Where robots in the early stages first did not have the capability of perception, they were mostly controlled by manual teaching, hard-coding and specific sensors (Hua et al., 2021). AI has the ability to transforms robots from mere programmable machines to dynamic entities capable of continuous learning and improvement.

TL can make the development of 'intelligent' robots even easier by allowing for a faster learning process. Transfer learning gives robots the ability to apply knowledge and skills acquired from previous tasks or environments to new, but related, situations. For example, a robot trained to pick up objects in a specific setting can transfer its learned skills to perform similar tasks in a different environment or with different objects, reducing the amount of new learning required (Hua et al., 2021). Alternatively, robots could be trained on data from simulated environments which is easier and cheaper to generate. Then, we a model is trained, the model can be transferred and applied to the real world. Research has shown that this is possible with the use of reinforcement learning (Z. Li et al., 2021).

2.3.5 Human Computer Interaction

The fifth domain in which TL could play a significant role is in better development of Human Computer Interaction (HCI) applications. HCI refers to how users (e.g. humans) interact with systems such as software applications, websites and robots. It relates to the research of how how these systems can adapt or personalize the experience for the users based on their behavior, needs, and preferences (Myers, 1998). TL in HCI mostly focuses on enhancing the user experience and interaction with software applications by enabling ML models to reuse knowledge learned in previous tasks.

For example, consider a recommender system used by an e-commerce website. A model trained to understand and predict user preferences and buying behavior in one domain, such as electronics, could utilize transfer learning to enhance its performance in a different but related domain, like home appliances. The pre-existing knowledge helps the model to quickly adapt and make relevant recommendations in the new domain (Fang, 2021). This enhances the interaction and the experience of the user by providing more relevant and personalized suggestions.

Another example could be the research by Nguyen et al. (2020) in which the researchers design a method to transfer emotion recognition of one system to another. This could potentially assist interactive chatbots with recognizing the emotion of the user they are speaking with better. This allows for a better understanding of the problem a customer might be dealing with and so assists customer better. TL makes it possible to significantly reduce the development and training time of these models.

Lastly, research has been successfully conducted that allowed researchers to develop better brain-computer interfaces. Brain-computer interfaces are tools which allow humans to send their brain signals to a computer by digitizing them. Then, those brain signals are classified and converted into actions. In this way for example, a human could control a computer simply by thinking about controlling it. However, one of the problems researchers face when creating ML models for this are individual differences between humans. Every human brain works a little different. By utilizing TL, researchers are trying to overcome these problems (H. He & Wu, 2019).

2.3.6 Reasoning

Next, the domain of reasoning comes across as a domain in which TL could have a positive impact. Reasoning for computers is described as the art of enabling computers to draw logical conclusions from data and knowledge (Anacleto et al., 2008). Additionally, computers should be able to make decisions based on those conclusions. Reasoning allows a computer to think for itself and to make logical decisions based on provided information.

Traditional AI reasoning faces challenges due to its reliance on hard-coded rules and knowledge bases, making it hard to adapt to new, unseen, or dynamic scenarios. TL acts as a mitigator to these challenges by leveraging previously acquired knowledge from similar domains or tasks. For instance, consider an AI model designed for legal reasoning to assess the validity of arguments in legal cases. If this model has been well-trained with a vast array of legal scenarios and cases, TL enables the application of this pre-existing knowledge in a related domain such as compliance and regulatory assessments in various industries.

By applying TL, the model doesn't start learning from scratch but uses its prior knowledge to draw logical conclusions and make decisions in new scenarios, making the reasoning process more efficient and robust. Such application of TL in reasoning helps in dealing with uncertainties, variabilities, and complexities by having a foundational knowledge base that can be fine-tuned and adapted according to the new reasoning requirements, leading to improved decision-making and problem-solving capabilities in AI systems.

Another example is the application of TL when making an AI play strategic games. With most strategic games, a good strategy is essential for winning. TL can help an AI in playing games by transferring knowledge gained about a winning strategy in one game to a different game. Current research has shown that a ML model specifically trained for a game 'PuckWorld' can be used as a basis for playing the game 'Snake'. By retraining specific layers they were able to boost the models performance in playing 'Snake' and made the learning process far more stable (Asawa et al., 2017).

2.4 Categorizations of TL

TL can be categorized in multiple ways. These categorizations all depend on the lens through which you look at TL. Literature has many ways of making a distinction in TL (L. Zhang & Gao, 2022) (Pan & Yang, 2009) (Zhuang et al., 2020), but generally speaking, the taxonomy of TL follows four main classifications, also shown in Figure 2.3.

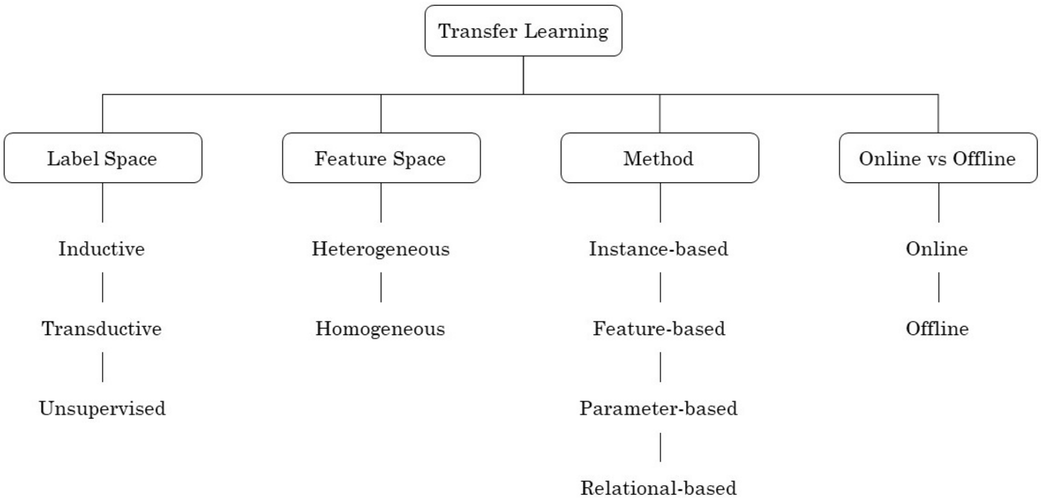


Figure 2.3: Classifications of Transfer Learning.

1. **Label space:** Do labels exist in the data of the source and target domain?
2. **Feature space:** Are the feature spaces the same between the source and target domain?
3. **Method**: Which learning strategy is being used?
4. **Online vs offline**: Is the learning strategy online or offline?

These are four questions someone can ask when looking at TL. All describe a different aspect of a TL process and distinguishes TL in a different way.

2.4.1 Label space

Based on the available labels in the data of both the source and the target domain, three categories of TL can be defined.

- **Inductive**: Inductive TL utilizes knowledge from a source task to improve learning in a related target task. It requires labeled data for the target domain. The source and target domains are required to be the same, but the source and target tasks the model is working on are different. The knowledge from the source model is then used to improve the target model. Inductive TL can be divided in two subcategories. When the source domain does not have labeled data, we distinguish this as Self-taught Learning, when the source domain does have labeled data, we distinguish this as Multi-task Learning.

- **Transductive**: Transductive TL focuses on transferring knowledge from a source domain to a different but related target domain, both dealing with the same task. Generally labeled data is here only available in the source domain, and not in the target domain. Instead of generalizing to new tasks, it generalizes across different distributions of data within the same task. For example, a model trained on photos taken during the day might be adapted using Transductive Transfer Learning to improve its performance on photos taken at night, without requiring night-time labels.

- **Unsupervised**: Unsupervised TL seeks to improve the learning of a target task in the absence of labeled data in both the source and target domains. The approach relies on uncovering and leveraging the intrinsic patterns, structures, or distributions shared between the source and target data. For instance, if a model is trained unsupervisedly to understand the structure of English news articles, it might be adapted to categorize French news articles effectively by identifying and utilizing inherent thematic and structural similarities, even when no labeled data is available in either language. This process facilitates knowledge transfer and model generalization across different, unlabeled datasets.

2.4.2 Feature space

A different approach to classifying TL is by the feature space. This was first introduced by Weiss et al. (2016). They classify TL in two main classes:

- **Homogeneous TL**: The source and target domains of a TL application have the same feature space.

- **Heterogeneous TL**: The Source and target domains of a TL application have different features spaces.

For example, in homogeneous transfer learning, you might have a model pre-trained on a vast dataset of English text for sentiment analysis, and then you adapt or "transfer" this model to analyze sentiment in product reviews, still in English. Both the source and target tasks are involving English text and sentiment analysis, making the transfer learning process homogeneous. On the contrary, heterogeneous transfer learning encompasses tasks that are quite distinct. An example of this would be a model initially trained to recognize objects in images. This model, after pre-training, could be adapted to a completely different domain, such as diagnosing diseases from medical images. In this case, the original and new tasks are different (object recognition versus medical diagnosis), manifesting the heterogeneous nature of the transfer learning process.

2.4.3 Method

This approach categorizes TL based on the techniques, algorithms, and methods used to implement and achieve a transfer of knowledge. The TL types are classified based on "what" part of the knowledge from the source is transferred (Curreri et al., 2021). Literature distinguishes four methods:

- **Instance-based TL**: Instance-based approaches are focused on reusing instances of the data from the source domain. By reweighing selected instances from the source domain, learning in the target domain could be enhanced.
- **Feature-based TL**: Feature-bases approaches are focused on transforming the feature representation, so that a positive transfer could be achieved. This approach can be divided in two sub-categorizations.
  - *Symmetric transformation*: Features from both the source and target tasked are transformed to align
  - *Asymmetric transformation*: Features from the source task are transformed to match the feature representation of the target task.
- **Parameter-based TL**: Parameter-based approaches are focused on reusing specific parameters from the source task, and applying these parameters for the target task. The parameters from the source model can be used as a starting point. By fine-tuning these parameters, a new target model can be trained.
- **Relational-based TL**: Relational-based approaches focuses on transferring the relationship among the data from the source and target domain. Instead of transferring only data-points, we transfer also the relational structures of the data. This is especially useful when working with data that has inherent relations, such as graphs, networks and relational databases.

2.4.4 Online vs Offline

Another method of categorizing TL is based on whether the learning process is being done online or offline. Currently, most TL applications are utilizing an offline scheme. In this approach, a model is pre-trained on a source task and later fine-tuned on a target task using a separate dataset, all in a non-continuous, batch manner. Conversely, in online TL, the model continuously learns and adapts to new data and tasks in real-time or near-real-time, making it more dynamic and responsive to evolving data landscapes and requirements (Zhao et al., 2014). Online TL is especially beneficial in scenarios where data streams are constantly changing and models require frequent updates to maintain performance and relevance.

Even the most used TL models (e.g. GPT, BERT) primarily operate based on offline transfer learning. The models are initially pre-trained on a massive corpus of text data, learning to predict the probability of a word given its context in a sentence. This pre-training happens offline and is a one-time process. After the pre-training phase, these models can be fine-tuned for specific tasks using additional datasets, which is also typically done offline. The fine-tuning adapts the pre-trained models to specialized tasks like text completion, translation, question-answering, and more, but this process does not happen in real-time or continuously online.

Online transfer learning could for example be a streaming service's recommendation engine, which initially learns from a dataset of user preferences from English-speaking regions. When the service expands it services to countries where different languages are spoken, the engine utilizes online learning to continuously adapt to new user data, including distinct cultural tastes and languages, in real time. By leveraging the pre-trained model and dynamically integrating the incoming data, the system maintains personalized content recommendations with increasing accuracy for a diverse, ever-growing user base, showcasing the model's ability to evolve with the changing data landscape.

2.5 From TL to Deep Transfer Learning

Deep Transfer Learning (DTL) is an extension of TL which specifically applies to Deep Neural Networks. Deep Learning is a subcategory of ML that utilizes NN's and is inspired by the way the human brain works. A NN consists of layers of nodes, an input layers, several hidden layers, and an output layer. Every node in the network is linked to another node and has an associated weight and threshold. If the output of a previous node is above the threshold of a node, the node is activated, passing data along to the next layer in the network (Han et al., 2018). The depth, or the number of these layers, distinguishes Deep Learning (DL) from traditional NN's. Deep learning models can have hundreds of layers. This enables them to process enormous amounts of data and automatically learn intricate patterns and representations.

DTL extends the concept of DL by using a trained NN as a starting point for a different but related task. The idea is to leverage the rich feature representations learned by the pre-trained model to benefit the new task, especially when there's limited data available for the latter. DTL is only related to the transfer knowledge with the use of NN's. It is given a special chapter however, because of the significance and amount of research that is being conducted on this topic.

The taxonomy and classifications of DTL in comparison to TL are mostly the same. However, DTL applies different techniques for model-based approaches (Iman et al., 2023). In general, we can classify these approaches as combinations of pre-training a model, freezing specific layers, fine-tuning specific layers, and adding extra layers. A Deep Learning Network that is trained on source-data and that can be utilized for TL tasks is called a pre-trained model. A pre-trained model consists of pre-trained layers with already initialized weights and parameters. This model can then be further trained on the target data. When layers are frozen, they are not altered in the re-training of the NN on target data. The weights and parameters stay the same and are simply reused. Fine-tuning layers means that the layers are altered during the re-training phase, but that the weights and parameters of the pre-trained model are used as a starting point and not randomly initialized. It is also possible to freeze a whole NN and add extra layers to a model to train it further on target data. This practice is also called Progressive Learning (Rusu et al., 2016). A schematic overview of an example of DTL is given in figure 2.4

The reason why this process of freezing, fine-tuning and adding layers works is the way in which NN's conduct the feature extractions in their layers. During the initial phases of training, a NN learns to identify various features and patterns in the input data across its multiple layers, progressively refining its understanding and representation of the data. Each layer of the network focuses on extracting different levels of features, starting from basic to more complex and abstract representations. Fine-tuning capitalizes on this pre-existing knowledge base, allowing for the adjustment and optimization of the network's weights and biases for a new, and

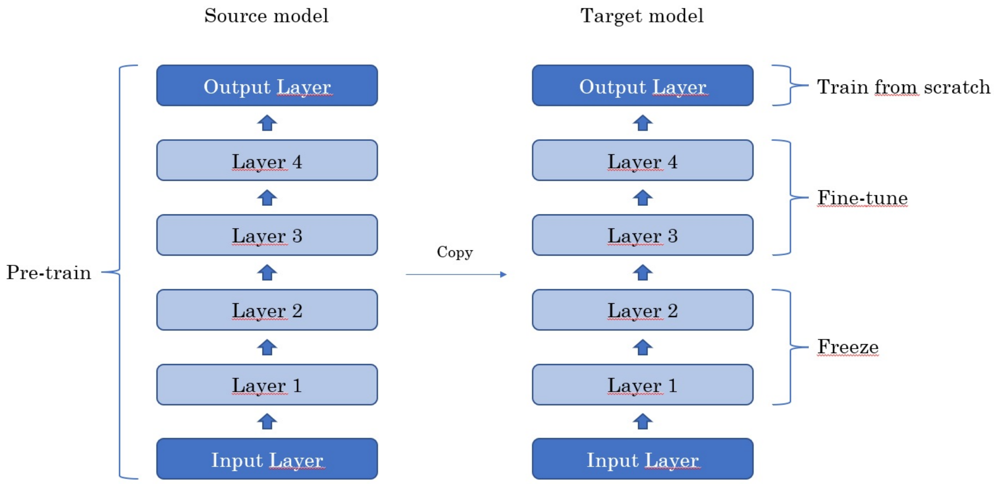


Figure 2.4: Overview of Deep Transfer Learning.

maybe more specific, task or dataset (Iman et al., 2023). It allows for a quicker and potentially more accurate adaptation of the network, leveraging the foundational learning previously achieved, and directing it towards improved performance on the new task. Thus, fine-tuning enhances the network's ability to generalize its learning to new, unseen data, promoting a more robust and adaptable model.

2.5.1 Adapter based DTL

Adapters in DTL represent another technique enhancing the adaptability and efficiency of NN's in transfer learning scenarios. Unlike more conventional methods which typically involve fine-tuning or freezing pre-existing layers of the network, adapters allow for the introduction of additional, smaller modules or layers into a pre-trained network. These adapter layers are then encapsulated between the pre-existing layers of the NN and are trained to adapt the model to the specifications of a new task, while the original weights and parameters of the network remain frozen (Houlsby et al., 2019).

This methodology provides a more flexible approach by enabling the model to maintain the robust, pre-learned features of the original network while allowing for customization to cater to unique characteristics of the new task. An advantage of using adapters lies in the capacity to allow task-specific customization without the need for extensive retraining of a network. This approach is not only computationally efficient but also mitigates the risk of overfitting, especially when the available dataset for the new task is limited.

When we compare adapter based TL and only retraining the encapsulated adapter layers compared to fine-tuning a full model, the generated results are very similar and the performance is almost identical. However, adapters are smaller in size compared to the full model and contain a significant smaller number of parameters to attain the same accuracy. This means that adapter based TL can provide a more energy and time saving friendly way of TL compared to 'normal' DTL.

In essence, adapters offer a balance, retaining the integrity and knowledge encapsulated within pre-trained models while providing a pathway to infuse them with the capability to adeptly handle new, diverse tasks. By doing so, adapters significantly contribute to enhancing the versatility and applicability of NNs in various DTL contexts, optimizing performance and fostering a nuanced adaptation process in alignment with the distinctive requirements of each task.

2.5.2 Foundation models

A relative new term for large pre-trained NN's is 'foundations models'. A foundation model is described as any model that is trained on broad data that can be adapted to a wide range of downstream tasks (Bommasani et al., 2021). Current examples include GPT-4 (Achiam et al., 2023), LLaMA (Touvron et al., 2023), and BERT (Zhou et al., 2023) for LLM purposes. Pre-trained models are not new, but models of this size which can with simple prompts to a good job in a wide array of tasks has called for a new classification. That is why the term 'foundation model' has been created. Foundation models allow fine-tuning on task-specific data which enables organizations and researchers to create more specialized AI applications more easily.

2.6 Transfer Learning in business

Thus far, we have focused on the different aspects regarding TL itself. Shown is that research regarding these topics is plentiful, and that TL provides solutions in all kinds of industries. In the following segment we will try to make the actual connection between TL and different business industries. How can organization reap the benefits of TL approaches? These industries are the healthcare, automotive and financial industry. Making this connection is especially valuable since TL is most of all a tool, developed to be used by organizations to help reach their goals. It is important to understand the value that TL could bring organizations across different industries.

In section 2.2 of this chapter we have identified five general reasons that could encourage organizations in the adoption of TL. In this section we will look at three different industries and identify theoretical case studies from scientific literature in which these general reasons can be identified.

2.6.1 Healthcare

The healthcare industry worldwide is utilizing more and more technology with successful outcomes. It is an important field for research that is closely related to everyone in the world. Better healthcare benefits us all. With the rise of AI, interests of applying these new technologies in the healthcare industry have also been rising.

In healthcare, deploying effective and adaptable ML models is often a challenge because of several barriers such as the scarcity of extensive, labeled datasets, stringent patient privacy regulations, not enough funding and resources, and the intrinsic heterogeneity and complexity of medical data. Traditional ML models, when applied directly, may not harness the full potential of available data due to these constraints.

Scarcity of labeled data can be a significant problem. One of the most common problems found is the scarcity of labeled medical image data. Research by Y. Zhang et al. (2020) shows how TL can be utilized in the Deep Domain Adaptation from Typical Pneumonia to COVID-19. Annotated COVID-19 data for pneumonia patients was scarce during the pandemic, but identifying and diagnosing pneumonia with computer assisted AI tools was life-saving. Thus, it was of great importance to develop ways in which ML models could be successfully trained on small amounts of labeled data. The researchers developed a model that was trained on publicly available pneumonia data and made it adaptable to images of people with COVID-19. TL was the preferred strategy in this use case.

By leveraging TL and specifically DTL, we can overcome more of these barriers. We can use lesser labeled data by utilizing large pre-trained models trained on data that is openly available. In this way, we need less specific labeled data for a specific task, which enables us to create personalized healthcare applications. A good example of this comes from an article by Chen et al. (2021). In their research a federated TL framework is suggested for wearable healthcare. Items like smartwatches and -phones are able to capture information of patients in an efficient way. With these data, we could develop models that provide early warnings to cognitive diseases or small vessel diseases which are personalized per patient. However, it is often a challenge to develop these models so that they fit people personally, since we need very specific models for this. TL could be a solution for this, the researchers suggest a framework based on federated learning and TL in which the privacy of the users is guaranteed, but the models can still be trained for personal use.

Federated transfer learning is also well-suited to the healthcare industry for specific privacy reasons because it allows for the development of reliable predictive models while keeping the used data, which is often sensitive personal data, localized. In traditional ML, data from various sources would need to be centralized to train a model. This raises significant privacy concerns and regulatory compliance issues with laws like the European GDPR. However, federated learning enables hospitals and clinics to train algorithms on their datasets without ever sharing the actual data. This decentralized approach means that personal health information remains on the local servers of the institution that collected it, with only the model itself, minus any individual data, being transferred and aggregated to enhance the model's performance (Topaloglu et al., 2021). By doing so, federated transfer learning upholds patient confidentiality, secures proprietary information, and facilitates a collaborative approach to improving healthcare outcomes across institutions without compromising privacy.

These examples show that TL has great potential in the healthcare industry based on scientific literature and proposed use cases. By applying TL techniques we could make the healthcare industry more efficient and more accessible to everyone while preserving privacy of patients.

2.6.2 Automotive

The automotive industry has always been considered as a leading industry regarding innovation. There are numerous amounts of technical opportunities and challenges. This is especially the case with the current shift towards electric and autonomous vehicles. If AI can play a central role here, so can TL.

The automotive industry is a broad industry where TL can be applied in all different kind of tasks. In this thesis not every aspect can be covered, but some examples of how TL can be used will be given. Accordingly, the focus will be on the most impactful areas where TL can significantly enhance performance and efficiency within automotive applications.

The majority of TL research in the automotive industry is focused on the development and improvement of autonomous driving systems. TL can help to be a useful tool for this in multiple ways. Generally speaking, developers can utilize TL to train algorithms that are important for tasks such as improving real-time object detection (Antunes et al., 2022), lane recognition, automatic braking, traffic lights recognition and weather condition recognition. Numerous scientific articles exist in which TL is used to develop these kind of systems. Take for example the work of J. Kim and Park (2017). In their research, the authors propose a sequential end-to-end TL method to estimate left and right ego lanes directly and separately without any postprocessing. Ego lanes are the lanes in which a vehicle is travelling. It is important for self-driving cars to correctly identify their ego lanes so they can stay in their lane which is crucial for safety and efficiency. The authors find that their TL approach improved accuracy and stability in comparison to traditional deep learning, while using lesser annotated data.

Another example that can be seen in the automotive industry are automatic traffic light recognition systems, which are developed for self-driving cars. These systems are important so that these cars can avoid collisions and are driving in a safe environment. Self-driving cars need to be able to identify traffic lights correctly and make a good decision on whether it is safe to drive through for example an intersection or not. These systems are of great importance for self-driving cars to avoid collisions and provide a secure driving environment. Autonomous vehicles should be able to correctly identify traffic lights and determine whether it is safe to drive or not. Research by Gautam and Kumar (2022) shows that TL based AI systems can achieve remarkable results and detect the correct traffic light situation with a very high accuracy. TL can be a much needed technique because it can significantly reduce the amount of needed training data, training time, computing costs, and everyday using cost of the model (Kheder & Mohammed, 2023). This can be especially important if we want these models to run efficiently in cars.

Another application of TL in the automotive industry, which shows that not everything is focused on autonomous driving, can be found in the development of predictive maintenance systems for cars. TL research on this topic is limited, but general ML research on predictive maintenance has been found. A survey by Theissler et al. (2021) shows that ML models have the potential to significantly ensure the functional safety of products in the automotive industry while keeping costs down. TL is discussed briefly in this report, the authors state that it is a promising field but that more research is needed. They do however cite some articles in which TL systems for predictive maintenance is proposed, but these are not specifically focused on the automotive industry. The general idea of TL in this sector could be that general algorithms trained on large datasets from a wide range of vehicle models and types can predict potential failures or maintenance needs. These models can then be fine-tuned for specific vehicle models, making them more efficient at predicting maintenance issues. Predictive maintenance supported by DL is already being integrated in the automotive industry, but TL has the potential to take this one step further. By not only personalizing TL models based on the vehicle type and model, but also on the driving style of the driver. Maintenance of cars could be made more efficient.

In conclusion, TL in the automotive industry is an interesting topic where more research is needed to identify all the current practical applications. It is in particular interesting to see in which way TL can make a contribution to autonomous driving and predictive maintenance, but more research is needed to show how organizations in this industry are actually implementing it.

2.6.3 Finance

The financial industry, a key driver of global economic systems, is also undergoing an almost mandatory transformation driven by the development of advanced technologies, particularly in the realm of AI. This transformation is not just a matter of technological advancement, but a fundamental shift in the approach to data analysis, risk management, and customer service.

In this context, implementing robust and efficient ML models is a critical challenge. The financial sector is characterized by its complex, high-dimensional data, stringent regulatory compliance requirements, and the need for models that can adapt to rapidly changing market conditions is growing. Traditional ML techniques often struggle to cope with these demands, particularly due to limitations in data accessibility, concerns over data privacy, and the dynamic nature of financial markets. TL could in this industry be a powerful strategy to address these challenges, offering a way to harness source data from domains and pre-trained models on large datasets and adapt them to specific financial tasks. This approach might not only enhances model performance but also reduces the time and resources required for model development, a crucial factor in the fast-paced financial environment. TL has the last couple of years definitely gained popularity in the field of finance and there are numerous applications for TL in this industry.

For instance, we can look at the application of TL for the risk assessment of credit. When financial organizations want to give out credit they may have limited or no data available on the historical lending outcomes of applicants, which makes it hard to develop accurate risk assessment models. This is in particular a problem for small and medium sized enterprises, because it makes it difficult for them to secure their needed financial support to grow their businesses (Suryanto et al., 2022). TL could be a solution because when data from different but related credit risk domains is available, we can transfer the knowledge and apply it in the source domain. As an example, in different research by M. Wang and Yang (2021) the authors propose a personal credit risk assessment model based on instance-based TL. Their results show a 24% increase in accuracy in comparison to traditional ML models.

A different application is time series forecasting with the help of TL. Time series forecasting involves the prediction of future financial trends based on historical data. This can include stock prices, market indices, interest rates, and economic indicators. Financial analysts use time series models to identify patterns and trends in past data, which help in making informed predictions about future market behaviors. This approach is important for investment strategies, risk management, and decision-making processes in finance. The accuracy of these forecasts is pivotal, as it directly impacts investment decisions and risk assessments in the highly dynamic financial market. Forecasting financial time series with the help of AI is getting more and more interesting the last years. Additionally, research shows that models based on TL can be superior in comparison to other base-line methods (Q.-Q. He et al., 2019).

The last application we will look at is focused on the utilization of NLP in the financial industry. The last years, there has been a growing interest in the mining of financial text and the automatic analysis of this. Monitoring markets through understanding financial communications is crucial for organizations that want to be successful in an online and rapidly moving market. Using NLP to efficiently read and understand these big corpa of text is an task which is being used more and more. However, until recently, NLP models were not specifically trained for the analysis of financial texts and statements. But, in the last years, researchers have created NLP systems with the help of TL that can do just this (X. Li et al., 2023). An important side note to make is that even larger models such as GPT-4 are challenging these smaller TL based models already. It does however show that TL can be of great use in the analysis of financial texts. Take for instance FinBert, a fine-tuned model based on BERT (Araci, 2019).

2.7 Gartner Hype Cycle

The Gartner Hype Cycle, a creation of the American research and IT advisory company Gartner, is a visual tool that illustrates the progression, acceptance, and societal implementation of various technologies. This model aims to graphically and conceptually depict the development stages of emerging technologies across five distinct phases. In this way, Gartner tries to provide clarity to organizations on the question whether a hype for a new innovation is justified (Gartner, 2023).

The X axis represents time and the Y axis represents the expectations. By moving from left to right, innovations are tracked as it moves through five predictable phases. These five phases that every innovation can go through according to the Gartner Hype Cycle are:

- **Innovation Trigger:** The first phase shows the emergence of a new technological innovation. Innovations in this phase generate significant media attention with promising proof-of-concept work and venture capitalists are starting to invest. However, often no usable products exist and commercial viability is unproven.

- **Peak of Inflated Expectations:** During the second phase, an innovation is at its peak of expectations, there are more and more users, but limited prove that the innovation can actually deliver. Early publicity has produced a number of success stories, often accompanied by scores of failures, leading to inflated expectations.

- **Through of Disillusionment:** When original excitement wears off and innovations fail to deliver, the expectations drop, entering the third phase. Technology producers only survive if they resolve issues and improve the performance of their innovations.

- **Slope of Enlightenment:** In the fourth phase, more instances of how the technology can benefit the industry emerge, leading to a better understanding and more practical applications.
- **Plateau of Productivity:** In the last phase, the innovation becomes increasingly stable and evolves into a more productive phase. Mainstream adoption starts to take off, which can be measured by benefits and wider usage.

2.7.1 Transfer Learning on the Hype Cycle

Gartner has not only invented the Hype Cycle, but also conducts research on where new innovations can be placed. According to the latest reports, they place TL in retrospect to Generative AI in the first phase, "Innovation Trigger". They predict that a Plateau of Productivity will be reached in 5-10 years (den Hamer, 2023). According to this report, we can say that TL probably still has lots of potential and it will take years before we know what the best applications for it are in organizations. Organizations are talking about it more and more and there is a beginning trend in capital that is being invested. TL is seen as an enablement technique of which we have not yet discovered its true potential. A visual representation of TL on the Hype Cycle is shown in figure 2.5.

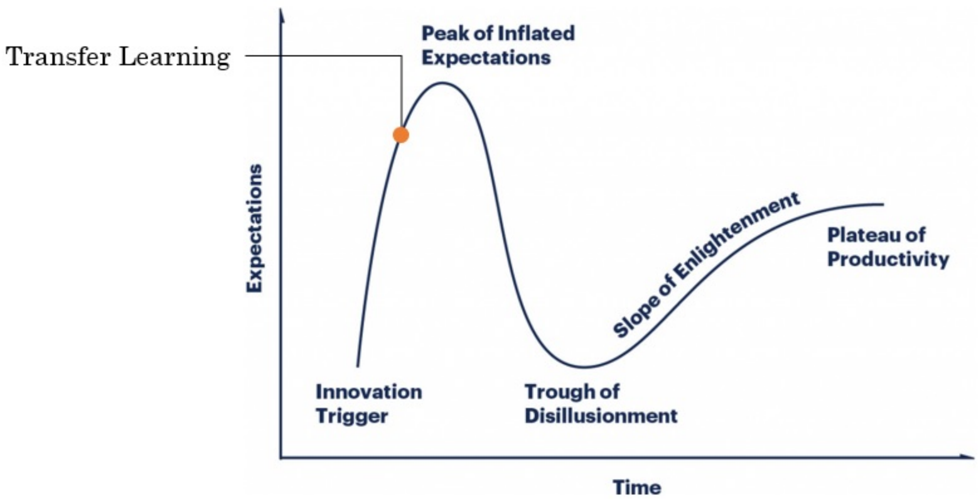


Figure 2.5: Transfer Learning on the Gartner Hype Cycle (den Hamer, 2023).

2.7.2 Criticism

It is important to note that there have been a number of critical assessments on the Hype Cycle. Some critics state that the phases are not scientific in nature and that there is no data or analysis that shows that the cycle is correct (Steinert & Leifer, 2010). Additionally, Gartner does not publicly disclose the specific methodologies used to plot innovations on the cycle. This lack of transparency can raise questions about the criteria and data sources used to assess the maturity and impact of innovations. Another critical point is that the Hype Cycle is technically not a cycle, it does not show that the events repeat in a circular fashion, but it is a linear progression through time.

Chapter 3

Methodology

The approach used in this thesis to address the research question will be explained in the following chapter. The research question is:

How is Transfer Learning currently being adopted across different industries and what are challenges and opportunities based on the identified TL approaches?

3.1 Research Strategy

The chosen research approach to find an answer for this question is a qualitative one. Qualitative research can be described as a series of interpretive methods that aim to explain, interpret, and translate ideas and occurrences rather than only document how frequently particular phenomena occur, as is often the research goal of quantitative studies (Basias & Pollalis, 2018). Qualitative research is less structured in comparison to other methods because it is focused on the formulation of new theoretical concepts (Williams et al., 2007). Additionally, qualitative research is often used to conduct research regarding the answering of experience and perspectives of people in a specific field (Hammarberg et al., 2016). Since the research in this thesis tries to bridge a gap between what is found in scientific literature and what is actually happening in organizations, a holistic view of the use of TL in different organizational domains is needed. A qualitative research design is seen as a best fit to identify detailed challenges and opportunities organizations face regarding the use of TL. To apply a theoretical framework of TL in different organizations, a deductive qualitative approach is chosen after developing an extensive theoretical framework.

3.2 Data Collection

Data has been collected by conducting 11 interviews with professionals. Interviewees were preferably employed in three specific industries; healthcare, automotive, and finance. However, due to challenges in finding participants within these industries, interviews were also conducted with professionals from other sectors. Despite not being directly involved in healthcare, automotive, or finance, these interviewees provided valuable insights into the application of TL in diverse organizational contexts.

Even though the research approach is deductive, an interview based approach has been chosen for practical reasons. It was expected that finding enough respondents for a survey who are proficient in TL would have been too big of a challenge for a thesis. The interviews were designed to gather insights and perspectives on the adoption, challenges, and opportunities associated with TL, specifically in comparison to the theoretical framework established by the preliminary literature review. In addition, an advantage of qualitative interviews in comparison to quantitative methods is that it allows for the recording of more detailed opinions of professionals.

3.2.1 Participants

Participants were selected based on their expertise and experience in implementation or working with TL. The intention was to speak to both technical experts and professionals specializing in more organizational aspects. The aim was to have an equal distribution between technical and organizational interviewees. This would ensure a broad understanding of the subject matter from multiple perspectives within the industries.

3.2.2 Interview structure

The conducted interviews were semi-structured. When gathering opinions on a specific subject, semi-structured interviews are the recommended method for gathering qualitative data. Specifically when the interviews serve as the bridge between literature and organizations. It also guarantees a certain level of consistency and it keeps the discussion focused on offering answers to the research question. This means that the interview questions have all been determined before the interviews, but that there is room for improvisation based on the given answers by the interviewees. This allowed for a good focus on the most important aspects on which the data collection was focused on.

The interview questions were structured in two parts:

1. **Introduction**: Interviews started off with introductory questions about the background and experience concerning TL and AI of the interviewee. The subject of the thesis was introduced and some general AI questions were asked.

2. **General TL questions**: General questions about TL and implementation of AI in one of the three industries and corresponding organizations were asked. The reasons for the use of TL as specified in section 2.2, challenges for the use of TL and the position of TL on the Gartner Hype cycle were discussed. Additionally, a link was made between the status of TL in the current world and in the future.

The full list of interview questions is included in the appendix under section A

3.2.3 Interview protocol

All interviews have been recorded with consent of the interviewee and have later been transcribed. To ensure a high quality of data and privacy for the interviewees, all interviews have also been anonymized. The industry in which the participants worked with TL was of great importance however and have been preserved, but the names of interviewees and organizations which were not relevant for the results in this research have been left out. The duration of each interview was expected to be a maximum of one hour. This time windows proved to be highly feasible. Interviews have been conducted both in-person and online via Microsoft Teams, whichever method was the easiest accessible to both parties.

3.3 Data Analysis

3.3.1 Transcription and data preparation

Initially, all recorded interviews have been be transcribed verbatim. This process converted the spoken words into written text, which was essential for a thorough analysis. During the later process of non-verbatim transcription, special attention was given to maintaining the accuracy of the transcripts while ensuring that the context and nuances of the responses were preserved. This was crucial because spoken word of respondents does not have the same sentence structure as written language; therefore, transcriptions of spoken words were adapted for clarity and readability, enabling easier analysis later in the process.

3.3.2 Coding and thematic analysis

Following transcription, the data has undergone a process of coding. When a transcription is coded, the text is categorized into segments and is assigned labels (codes). These codes have been written down in an iterative manner. They started from a set of preliminary codes based on the interview questions and have been expanded to include new codes that emerged from the interviews. The coding process consisted of three phases. These phases are based on the Grounded Theory process of coding (Thornberg, Charmaz, et al., 2014).

1. **Open Coding**: Open coding is an initial phase where the data is broken down analytically. In this stage, interview transcripts are read thoroughly, and initial codes are assigned to describe the content. These codes are often descriptive and are derived directly from the data. Open coding allows for an identification of significant categories, patterns, and themes which emerge from the data without any preconceived notions.

2. **Axial coding**: Following open coding, axial coding is undertaken. This stage involves connecting the codes identified during open coding and grouping them into broader categories. Axial coding helps in identifying relationships between different categories and subcategories. It serves to organize the data more systematically, facilitating the identification of main themes and patterns.

3. **Selective coding**: The final stage of coding is selective coding, where the focus shifts to integrating and refining the categories into a coherent story. In this stage, core categories and central themes which represent the main findings

of the research are defined. Selective coding involves relating all other categories to this core category, thus synthesizing the data into a comprehensive story that addresses the research question.

After coding, the next step that has been performed is a thematic analysis. This involved the identification of patterns and themes within the codes. Themes are overarching ideas or trends that are prevalent across multiple interviews. This analysis has been guided both by the research question and the theoretical framework established in the literature review. The aim was to identify commonalities and differences in the adoption, challenges, and opportunities of TL across the healthcare, automotive, and finance industry.

3.3.3 Interpretation and linkage to theoretical framework

The identified themes have then been interpreted in the context of existing literature on Transfer Learning. This step was critical for bridging the gap between theoretical concepts and practical implementation of TL in different industries. The interpretations have been focused on how the insights from professionals aligned or diverged from the established theories that have been derived from the conducted literature review.

3.4 Reliability, validity and ethical considerations

When conducting qualitative research, it is important to ensure a high reliability and validity. This establishes the trustworthiness of a study's findings. In the following section, the strategies that were taken to ensure this have been described. Additionally, the ethical considerations that were taken are also described.

3.4.1 Reliability

With reliability in qualitative research, we look at how consistent the research method is in it's measures. In quantitative research, reliability refers to the possibility to exactly replicate the results of a study when you conduct it again in the same environment . This however is more difficult in a qualitative research approach when conducting a small amount of interviews. This is why consistency in your research approach is considered on of the most important factors for reliability in qualitative research (Leung, 2015). Based on this, some measures have been taken to ensure a high reliability.

Standardized Interview Protocol: A consistent interview format and set of questions are used across all interviews. This standardization helps in maintaining a consistent data collection.

Clear interpretation: Findings have been documented in a clear and transparent matter. During the whole process of interviewing, important actions and steps such as interview procedures, coding decisions and thematic analyses have been written down. This ensures that the process can be repeated.

3.4.2 Validity

In addition to striving for a high reliability, it is also important to strive for a high validity of your research's results. Validity in qualitative research refers to the accuracy and truthfulness of the findings. In other words: how true your findings are. This can be difficult to measure and this is why validity is can also be described as the 'appropriateness' of your tools, processes and data (Leung, 2015). This refers to how valid your research question is for the desired outcome, whether the research design is valid for the methodology and whether your sampling and analysis is appropriate in the context of your research. To ensure a high validity, two often described measure in scientific literature has been taken, which is described below.

Purposive sampling: Since the research subject is highly technical, a random selection of participants could lead to disappointing results. Because of this, a purposive sampling technique has been chosen. A purposive sampling technique is a non-random sampling technique where participants are selected based on specific criteria that are relevant to the research objectives, in this case hands-on experience with TL (Robinson, 2014).

Information saturation: When no new information or themes are emerging from conducting more interviews, a state of information saturation has been reached. Achieving information saturation is often considered an indicator of the completeness and depth of data collection (Saunders et al., 2018). To meet this requirement, multiple interviews were conducted per industry.

It is important to note that the author is aware of more strategies to ensure high validity in qualitative research. Some of which are triangulation and participant validation. Unfortunately however, because this research is a solo project with a deadline and limited resources, these measures could not be taken.

3.4.3 Ethical Considerations

Ethical considerations in qualitative research are described as a set of principles to ensure that your research is conducted in a fair and ethical manner. Especially in this technologically oriented world, it is important to consider ethical considerations such as informed consent and privacy of your respondents (Eysenbach & Till, 2001). This section describes the measures that have been taken for this.

Informed consent: Before the interview started, participants were informed of what the research was about, what their role was in this research and what possible implications were of them joining this research. After that every participant was asked for consent of joining the research. This consent was given verbally.

Anonymity: Respondents were ensured that participation in this research was anonymous and that it would not be possible to disclose their own or their employers' identity. This was done by anonymizing every interview. It was important to do this because it allows participants to talk freely without fear of personal identification or repercussions.

Data security: To protect research data from unauthorized access or breaches all data has been securely stored. Ensuring data security is crucial to maintaining participant confidentiality and upholding ethical standards.

Transparency: It was important to be as transparent as possible not only participants to participants, but to everyone who has been involved in this study. Clear communication and openness throughout the research process was therefore always held in high regard.

Chapter 4

Results

In the following chapter the results from the interviews will be presented.

4.1 Demographics

11 Participants were interviewed in total. All of them had experience with TL in their field, their view on this however differed. All participants were considered at least senior level in their line of work. Some participants had a more business oriented background and some had a more technical oriented background. Table 4.1 gives an overview of the conducted interviews and shows the background, the current field of work and the current function of the participant. Please note that participant F2 has experience in both the finance and the healthcare industry. However, since their main position was currently in finance, they are grouped under the finance (F) participant code.

Four Participants who are classified as 'General' were not employed in the three main industries, but did have hands-on experience with or knowledge of TL. Nonetheless, their interviews gave a welcoming different view on TL as a whole. Participant G1 had experience in the Agro-industry. Participant G2 was currently employed in audit and assurance. Participant G3 worked as a data scientist for all kinds of fields, including the public sector and transport. Finally, participant G4 was employed in the recruitment industry.

Participant code	Background	Field of work	Function
A1	Technical	Automotive	AI Lead
A2	Technical	Automotive	Technical Manager
H1	Business	Healthcare	IT Consultant
H2	Technical	Healthcare	AI Researcher
H3	Business	Healthcare	AI Product Specialist
F1	Technical	Finance	Data Scientist
F2	Technical	Finance & Healthcare	Head of Data and AI
G1	Technical	General	Software Engineer
G2	Technical	General	AI & Ethics Lead
G3	Technical	General	Data Scientist
G4	Technical	General	AI Lead

Table 4.1: Overview interviews

4.2 Findings

4.2.1 Current applications and future prospects of TL

Multiple specific applications of TL which have been identified in the three industries, together with various more general applications not bound to a specific industry are presented in this section. Alongside this, multiple respondents spoke about the AI and TL strategies of their organizations and what their views are on this in the future.

Automotive

The first industry that was researched was the automotive industry. Two respondents have been interviewed who were currently employed in this field. They reported multiple current use cases of TL and also discussed potential future applications.

One of the respondents was currently employed by a major car manufacturer. In his work he specialized in the development of AI applications for motorsports. He gave two examples of how they currently used TL in their work.

Firstly, during testing, race cars would be equipped with numerous sensors to analyze their racing performances. However, due to racing regulations, these sensors must be removed when participants engage in races. To address this, the respondent stated that they used AI to simulate the sensors. The sensors are trained using time-series data from the testing phase. A challenge arises when the aim was to test on as many racing tracks as possible, which was not always feasible. To combat this, when racing on unfamiliar tracks, they would use the same ML model used in other tracks as a starting point and fine-tune that with a minimal amount of new data (ParticipantA1, 2023). Their solution was to apply TL.

> *"The new network must adapt quickly with minimal new data. We achieve this by utilizing pre-training on a network whose domain is quite similar. While the sensor may not be entirely different, it might involve a slightly varied sensor, and we incorporate newly collected data to retrain the sensor (ParticipantA1, 2023)."*

Another example, also related to motorsports was how they used TL to detect objects in front of their race car. They would use pre-trained NN's for general object detection and fine-tuned them to specifically identify race cars (ParticipantA1, 2023). Fine-tuning was essential because image data of race cars could significantly differ from original data the pre-trained NN's were trained on. They only retrained the final layers of a NN for this application.

Additionally, applications for TL that lay outside of motorsports were also discussed. One respondent used TL for their smart-manufacturing purposes. Both respondents from the automotive sector confirmed that this was a viable application for TL. In the context of car production, TL can be used to monitor cars when they come off the production line. Computer vision solutions can check if all the parts are on the car and if everything is installed correctly. Or it could check if certain parts are produced correctly on the assembly line. This was considered very important since it often happens that parts are missing. One respondent stated that almost all computer vision application they developed were based on TL. They would start with for example a ResNet model trained on ImageNet and fine-tune it to be able to identify car parts (ParticipantA2, 2024).

> *"Smart manufacturing, you can imagine tasks of checking if a part on a factory line is good or not, for instance. You train on good and bad examples, and then the model will give you whether a specific part is correct or not, is well-manufactured or not (ParticipantA2, 2024).*

TL could be used in the context of manufacturing because every car is a little different. You could start with a general model and fine-tune it to fit specific car types.

> *"You would essentially apply transfer learning from, let's say, all the cars in your fleet. Train the model with existing data and then adapt it for new cars. The pre-trained network can be used to accommodate a slightly different domain, like transitioning from a regular small car to a pickup truck, and so on (ParticipantA1, 2023)."*

Other possible use-cases of TL in the automotive sector, with which the respondents did not have hands-on experience with were also discussed. One respondent noted that TL could potentially be a major deal breaker for real-time object detection in autonomous vehicles. Currently, organizations which excel in manufacturing those cars, are the ones who have the ability to collect the most amount of data, this gives them a major advantage for the development of high quality AI models. However, since TL facilitates that high quality models can be shared, it could transform the sector. Not every company has to build models from scratch themselves, but a market of models might become available, potentially democratizing AI development.

> *"Models that can be easily adapted will gain inherent value and can be traded. This trading aspect makes the transfer learning approach particularly intriguing (ParticipantA1, 2023)."*

The respondent stated that this might even facilitate the application of specific models trained for one sector, in the application of a different sector. Adapting the same models across diverse contexts could provide interesting possibilities.

> *"These models, however, might find applications in entirely different domains, such as Netflix. For instance, using object detection in movies to identify a BMW and then showing a commercial on Netflix about BMWs (ParticipantA1, 2023)."*

With the rise of LLM's and foundation models, respondents also noted that TL could be fruitful in this sector of AI. One respondent stated that he believed all NLP solutions in organizations are TL based (ParticipantA2, 2024). He explained that there is no reason for most organizations to develop NLP models themselves, because the benefits would not outweigh the costs.

> *"For large language models, I do have some experience, and then I think transfer learning is even more used. If some of the other companies you spoke to said that they do LLMs from scratch, they would lie (ParticipantA2, 2024).*

A specific example that was given was the fine-tuning of LLM's for digitized car manuals. By fine-tuning a LLM you could potentially create a chatbot like version or a version accessible through speech. This could then be integrated into the car. By starting with one general LLM for car manuals, TL could be used to fine-tune networks so that for every type of car a digital manual could be created (ParticipantA1, 2023).

Another application of TL could be the use of time-series data. One respondent noted that models which analyze how customers behave with or use their cars in specific regions or countries can be developed. By using a model that is concentrated on regions with an abundance of data, for instance Europe, as a starting point, that model can be adapted to fit different continents with less available data.

> *"Through pre-training and fine-tuning, you adapt the network to a slightly different domain, which may involve different continents, behaviors, or types of cars sold there (ParticipantA1, 2023)."*

Healthcare

The healthcare industry is a big industry which focuses not only on practical applications, but also heavily on research. Overall, four main types of AI applications have been identified, all in which TL can be used as a viable strategy. Additionally, for all AI types, actual TL applications in practice have been identified.

The first category of AI encompasses a broad range of classification and segmentation challenges. These are mostly based on image forming research. Application consist of actions that people could in principle do just fine if they are given enough time for it. However, often vast amounts of data have to be processed to find relative small anomalies. For instance, big amounts of patients that need to be classified in a high or low priority; triage problems. AI can then be a more efficient solution.

Multiple TL applications have been identified for this AI type in the healthcare industry. Respondents stated that TL for classification and segmentation issues is a widespread technique and that the technology in this field is not new. One respondent gave multiple examples after he was asked how he used TL in his work.

> *"Think of segmenting a stroke in deceased animals, finding thrombosis, or segmenting brain tissue to determine volumes. And in your heart, for example, you can look at what the ejection fraction of blood is that is being pumped out. Let's say, a CTA-like situation (ParticipantH2, 2024)."*

A different respondent noted that TL is being used in image classification applications. He stated that models can be adapted and used for a different task in a relative simple manner.

> *"But usually, a model capable of detecting a certain anomaly in an image can also do the same with relatively little additional data in a different setting. I know that such things are already being applied. So, for example, there is an examination of certain forms of anomalies in imaging studies (ParticipantH3, 2024)."*

The second type of AI in healthcare is the use of predictive models that look at which patients could potentially develop complications or which patients are having a higher chance on having a specific condition. Based on these models, the healthcare industry for example tries to predict which medications should be given to patients (ParticipantH3, 2024). Types of AI that are developed here are often more directed to the day-to-day medical practices of the field.

Respondents noted that TL is definitely beneficiary for this AI type. Especially because TL can accommodate more personalized AI applications. A given example was how personalized models made with TL had a smaller training time and lesser data need (ParticipantF2, 2024). However, multiple respondents stated that in practice the application of TL is still quite limited (ParticipantH1, 2024), (ParticipantH2, 2024), (ParticipantF2, 2024).

> *"The use of transfer learning is still quite limited in practice. At [current employer], in collaboration with [hospital], we have developed our own model. We are currently investigating how we can adapt this model to implement it in other hospitals as well (ParticipantH1, 2024)."*

Thirdly, there is the new development of the application of LLM's in healthcare. AI applications that fall under this are among other things focused on lowering the administrative burden of healthcare workers. For instance the automatic generation of texts and dismissal letters, the automatic conversion of medical terminology into layman's terms, and more simple but labor intensive tasks that you would rather not have done by a doctor or nurse.

TL, and specifically the practice of adapting and fine-tuning foundation models, is a technique that is gaining in popularity for this AI type. All respondents stated that fine-tuning is a promising technique and they expect more applications for the healthcare industry in the upcoming years. One respondent noted that she had recently witnessed a presentation of how fine-tuned LLM's would be rolled out in a hospital she worked for (ParticipantH1, 2024). Another respondent stated that they were developing a chatbot for their customers by fine-tuning a LLM on their internal data.

> *"We receive a lot of messages regarding support calls for our services and software. We have a whole team of people answering it all. But very often the same questions are asked. So what we actually did is just built a kind of private instance based on ChatGPT. With all previously submitted and answered support calls as training data (ParticipantH3, 2024)."*

The last AI application in the healthcare industry focuses on applications that not only lie in direct medical applications, but more on adding value in business operations of hospitals. One respondent stated for example that applications such as capacity planning of employees and operation room planning are being used (ParticipantH1, 2024)".

TL is also a viable method for this AI type. Respondents stated that models could be transferred between hospitals or that planning models made for specific employee types could be altered to fit others.

> *"So that's actually not specific to healthcare, but really planning. And there you see a very big added value (ParticipantH1, 2024).*

Next to the identified AI types, a special notion for Federated TL solutions needs to be given. Multiple respondents saw Federated TL in the healthcare industry as a viable solution, even tough it would be a complex operation. One respondent noted that Federated TL could be valuable to hospitals because it would allow them to share models without sharing direct privacy sensitive information. However, it needed more research and no actual implementation was yet completed (ParticipantH1, 2024).

Finance

The finance industry is an industry driven by technological innovation. AI is one of the technologies in which a considerable amount of resources is spent to enhance the sector in all sorts of ways. Multiple applications of TL in this industry have been identified.

Just like in the healthcare industry, a major role for the fine-tuning of foundation models has been identified, primarily focused on LLM's. The attention that is given to text analysis and generation with the help of TL is growing. Respondents gave two examples of how that was currently being developed or applied.

The first respondent worked for the HR team of a large bank and was together with his team tasked to create a model for sentiment analysis of customer chat data. This data was to be processed in batches. They first started to experiment with rule-based and lexicon approaches. This was successful. However, they later transitioned to classical ML algorithm approaches such as XGBoost, aiming to further enhance their results. When this did not satisfy their requirements, they started using a TL approach by fine-tuning the pre-trained BERT model. With this, they were able to get the results they needed while simultaneously reducing their workload (ParticipantF1, 2024).

> *"This batch processing approach meant that we could efficiently manage the inference time without significant delays. So, those were the key considerations in adopting BERT for our sentiment analysis, and overall, the outcomes were positive (ParticipantF1, 2024)."*

The second respondent, also employed by a large bank, gave another example of how they used transfer learning by fine-tuning foundation models. Their applications were more focused on Generative AI. The respondent could not give too many details, but assured that they used the fine-tuning of foundation models for, among other things, analysing data and automating specific pipelines or onboarding flows. The respondent noted that TL was often their preferred technique.

> *"Yes, generative AI itself is actually transfer learning. We use the same foundational models across different departments. And then we do fine-tuning or up-training of them per use case (ParticipantF2, 2024)."*

Additionally, TL applications have also been identified to be helpful for fraud detection strategies of financial institutions. Both of the financial oriented respondents agreed that TL can be a fruitful strategy. One respondent stated that TL can be used in fraud detection, in particular while considering their plans for further international expansion. He explained that strategy would be to develop a base model in one country, and then retrain and fine-tune it to suit different countries (ParticipantF2, 2024).

> *"Yes, for example, you could train the fraud detection model in the Netherlands and use it in the U.S. Because we are also working on the U.S. branch (ParticipantF2, 2024)."*

Another respondent noted that TL would be a good strategy for fraud detection because it accommodates 'online learning'. This relates to the online TL subcategory explained in section 2.4.4. The respondent stated that fraud detection models have the need to be able to quickly adapt to changes when there is not a lot of time to retrain a full model. This is especially the case for financial institutions, since they are processing vast amounts of data in real-time, such as online transactions (ParticipantF1, 2024).

> *"Another facet of transfer learning in ML is termed "online learning." This approach is not specific to text data and is commonly used in scenarios where the nature of the domain, such as fraud detection, necessitates rapid adaptation to changes (ParticipantF1, 2024)."*

The last given example by respondents was the use of TL in compliance and transaction monitoring, in particular for retraining models based on different types of subject data. The example that was given was that developers would start with a compliance model based on data of their retail customers. If they then wanted to develop a compliance model for a different user group, for instance business customers, they could use the retail model as a base and fine-tune it to fit business customers (ParticipantF2, 2024).

> *"So, for example, I can say, okay, I train a model for this group of users. These could be retail customers, for instance. And then you can say, okay, but for businesses, I start with the model that I had used for retail customers. And then I'll continue fine-tuning, so to speak(ParticipantF2, 2024)."*

The respondents also stated that even though TL was a viable strategy for them to use, it was often not needed. Organizations in this industry often have enough resources at their disposal to develop models from scratch (ParticipantF1, 2024), (ParticipantF2, 2024). One respondent said that he expected a bigger role for TL since the rise of foundation models, but traditional TL was given less attention (ParticipantF2, 2024).

Modelshops / model database

Several respondents not only discussed their current applications of TL but also shared their visions for the TL landscape as a whole. A reoccurring theme in this was the creation of modelshops for TL or pre-trained model databases. 6 out of 11 respondents discussed this as a big opportunity for TL. Platforms such as Huggingface that act as an easily accessible modelshop already exist, but participants specifically mentioned that they envisioned this internally for organizations as well. Organizations can build a database of pre-trained models together with a platform that allows for easy fine-tuning. This could allows organizations to quickly reap the benefits TL provides. Additionally, they predicted that the future of AI application in organizations would consist for a big part of TL. With some big organizations offering commercial pre-trained AI solutions which can be implemented in an easy way. Together with an active open-source community (ParticipantG2, 2023), (ParticipantG3, 2023), (ParticipantA2, 2024) & (ParticipantF2, 2024).

> *"Overall, there will be big companies that give you everything already ready and packaged to do your own things. There will be still a really active open-source community that will keep pushing things out that with a little bit of effort, you could use that are cheap to use. But then people that really want to do things without thinking about it much will go for commercial solutions. And that's the playground (ParticipantA2, 2024)."*

One respondent noted that they are currently in the process of envisioning and developing a modelshop for the healthcare industry in the Netherlands. However, he said that they should watch out for an abundance of specific models. Good model enforcement and maintenance would then be important.

> *"But our vision is also really about having a national or at least a more widely available AI library. Essentially, just like an app store. The tricky*

part about that is, I already see the risk that this app store will mainly be filled with 47,000 versions of a commercially developed complication predictor (ParticipantH3, 2024)."

Another participant, working in a different industry, specifically in recruitment, mentioned that they are actively developing a modelshop for internal use. This modelshop would allow engineers in other department to efficiently build more customizable ML models by utilizing TL. The models in the modelshop could either be trained from scratch in-house, or could come from an external source. He noted that this solution was not limited to their own industry and could also be used in for example the automotive, healthcare and finance sector.

> *"So we actually have a list of supported models. These will be centrally prepared and made available for use by other projects within the organization. Then they can build their applications on top of them (ParticipantG4, 2024)."*

4.2.2 Primary data types used in TL

Multiple data types that organizations use while applying TL have been identified. Respondents from all industries stated that they used both structured, unstructured and multimodal data for their AI applications. These data would also be suitable for TL purposes.

Structured and unstructured data

One respondent stated that the world of healthcare is increasingly making the transfer from the use of unstructured data to structured data. He stated that even though doctors often still prefer the usage of handwritten notes or text input fields to capture patient data, a quick transfer to a more structured way of data storage is being detected. Mostly in the form of coding-tables that are being stored in the Electronic Health Record (EHR) system. This allows the industry to more easily analyze their data with for example TL.

> *"So, coding tables are used. So SNOMED, ICD-10, a diagnosis thesaurus for example, DBCs. Everything is mapped as much as possible to coding systems, structured data. And all of that is simply stored in the database of the EHR (ParticipantH3, 2024)."*

Another respondent from the finance industry stated that they too use both structured and unstructured data for their ML purposes. Both of them would be suitable for TL. Structured data in their industry mostly consisted of tabular numbers and categories. Unstructured data mostly consisted of text, suitable for for example the fine-tuning of LLM's and foundation models (ParticipantF2, 2024).

Text data

All respondents stated that TL with the use of foundation models was for the majority based on unstructured text data. Foundation models were mostly discussed in the form of LLM's. Types of text data that have been identified in the three main industries are customer chat data, car manual data, patient data and customer feedback data (ParticipantA1, 2023), (ParticipantH3, 2024), (ParticipantF2, 2024). Outside of these industries one respondent stated that they mostly used resume and employee oriented text data for their TL purposes (ParticipantG4, 2024).

Image data

Almost every respondent noted that multimodal data such as images deserves a lot of attention in both traditional ML as in TL. Both the automotive, healthcare and finance industry stated that they use image data for their TL purposes. In the automotive sector examples of use cases were object identification on the road and smart-manufacturing controlling. The healthcare sector uses image data for radiology and the finance sector for the face-recognition access in their applications.

Outside of the main industries, image data was also widely used for TL. One respondent stated that they developed a TL application for the recognition of pears for the Agro-industry (ParticipantG1, 2023). Another stated that for a children's application in the zoo, a TL model was developed for the recognition of animals (ParticipantG2, 2023). A third respondent stated that they used TL for a model that detects earthquake related cracks in walls (ParticipantG3, 2023). They started with a model used for the detection of potholes and fine-tuned it to be applicable to houses.

Time-series data

The use of time-series data for TL has been identified in two industries. One of which is the financial industry. By utilizing time-series data a respondent stated that they created a feature in their banking app that predicts the spending habits of a user. This helps determine if someone has enough money to last until the end of the month (ParticipantF2, 2024). Although they did not use TL in this use case, the respondent stated that it would have been a viable solution.

Another respondent from the automotive industry stated that they used time-series data for their racing purposes. He stated that the use of time-series data for TL was more challenging compared to different data types. Image, text and speech data contain a finite amount of information which can more easily be leveraged in various layers of a NN. Time-series data was perceived to be more complicated because it is missing this structure. This makes fine-tuning a network based on time-series data more difficult.

> *"Handling time series data becomes more intricate because, without a clear structure, it becomes challenging to integrate it into various layers of your network. Consequently, freezing the initial layers, which capture the basic structure of the time series, and then retraining only the last layers becomes more difficult (ParticipantA1, 2023)."*

The respondent from the finance industry however disagreed with this statement. He stated that the use of time-series data specifically for TL is not more difficult. This shows a contrasting view on the use of time-series data. Nevertheless, ML application based on time-series data in general are harder than structured data (ParticipantF2, 2024). But he did not have a problem with TL for time-series data specifically.

> *"I would say no. The issue isn't transfer learning itself. The problem lies with the data. Handling time series with different modeling approaches is also challenging. That's not because of transfer learning. It's usually the data that isn't quite good enough for those types of tasks. When you combine it with transfer learning, then it indeed becomes challenging. But I don't think that has much to do with transfer learning (ParticipantF2, 2024)."*

4.2.3 Reasons for TL application in organizations

All reasons previously stated in the theoretical framework for the use of TL in organizations have subsequently been been identified among the respondents. Every reason has been ranked based on the personal experience of the respondent within their industry. This section presents all given answers of the respondents together with an explanation of how and why these reason are ranked in general. Additionally, some newly identified reasons are presented. For every identified reason, respondents have furthermore given their predictions for whether that reason would become more or less important for the application of TL in the future. Table 4.2 shows the average rankings of all respondents. All the individual rankings can be accessed in the appendix under section B.

Reason	Average ranking	Average expectation
Data scarcity	1.2	-
Lack of computational power	2.6	-
Maintaining privacy	3.6	+
Lesser environmental impact	3.7	+
Personalized AI	3.8	+

Table 4.2: Average rankings initial TL reasons

Data Scarcity

Data scarcity was mentioned as the number one reason for the current use of TL in organizations with the exception of two respondent. One placed data scarcity as number two, as seen in table B.6. He did recognize that data scarcity was a major incentive for the use of TL overall, however, data scarcity was not really an issue in their specific organization (ParticipantF2, 2024). Another respondent ranked 'Personalized AI' as number one, he reasoned that all TL solutions come from a need for some form of personalized AI (ParticipantA2, 2024). He ranked data scarcity as number two, as seen in table B.2.

Thus, respondents from the automotive industry both ranked data scarcity as one of the highest incentives for TL. One respondent stated that because of racing rules, they were limited in their data collection. TL was their solution to solve this data scarcity issue. Another example that was given was that data regulations are becoming stricter. It is getting more difficult for organizations to gather certain amounts of data. By using pre-trained NN's and applying fine-tuning strategies, they could overcome this problem (ParticipantA1, 2023).

> *"The other thing is that, specifically in European markets, but also in other markets like the US markets, data regulations are getting much stricter. We are basically hindered in collecting certain amount of data. We are just not able or allowed to collect data, and by using pre-trained neural networks, we are able to overcome that because we need much less data (ParticipantA1, 2023)."*

The second automotive respondent stated that all their AI applications were based on TL. Data scarcity was a very common problem (ParticipantA2, 2024).

In the healthcare industry, the number one incentive for TL was also data scarcity. Multiple reasons for why this was the case have been identified. One respondent stated that gathering enough data for ML purposes is often a challenge. The healthcare industry is almost always in a shortage of data (ParticipantH3, 2024). For example, training models to identify exceptionally rare diseases can be very difficult because so little data is available. TL could then be a solution.

> *"Sometimes, for example, it's not even available within the Netherlands. Especially when it comes to relatively rare conditions, for instance. Then we notice that it's already difficult just to find some data, let alone enough data to truly have a representative representation (ParticipantH3, 2024)."*

> *"One reason could be that rare conditions, such as Madelung syndrome, characterized by abnormal growth in the arm, are infrequent. In such cases, data availability is limited due to the rarity of the condition (ParticipantH2, 2024)."*

A different reason for data scarcity in the healthcare industry is not that data is scarce, but that the access to the data is limited. A respondent noted that often enough data about specific conditions might be available in the country as a whole, but that because of privacy regulations they could not access these data. Then, TL, or a federated TL framework would be an ideal solution. Another respondent said that in for example the U.S., privatized hospitals are not willing to share their data, which in turn also limits AI capabilities of the industry.

> *"So, what I actually mean is that although the data is available, the amount of data in one hospital is often too limited to be effective. That's why transfer learning is so valuable because it enables hospitals to benefit from larger and more diverse datasets, even if their own data is limited (ParticipantH1, 2024)."*

> *"In the U.S., hospitals sometimes do not want to share their data, leading to isolated data sources (ParticipantH2, 2024)."*

Respondents employed in the financial industry in turn also confirm data scarcity as the primary reason for organizations to apply TL. One respondent stated that data scarcity can occur in between different departments of financial organizations. Every department or section of an organization normally has a data owner and data is not allowed to be freely shared among departments. This has been done so that the financial organization of the respondent could comply with laws and regulations. He however stated that it can inhibit ML applications, since data might not be accessible to certain development teams. Collecting and centralizing these data might take years to gain approval from various stakeholders (ParticipantF1, 2024). When that is the case, TL could be used to overcome these issues.

> *"Another factor to consider, especially in larger organizations, is that each section has its own data owner. This means that data isn't freely shared among all departments, and the use of data is highly regulated (ParticipantF1, 2024)."*

Other general respondents stated that current reasons for data scarcity included for instance the ever growing need for more data. We are collecting more and more data, but never have enough for very specific use cases. That is why data scarcity would always be a reason for TL (ParticipantG3, 2023). Another respondent stated that when because of time and money constraints for a project they could not gather enough data, they would apply TL (ParticipantG2, 2023).

When asked about how data scarcity would influence the choice for TL in the future, the general consensus among respondents was that data scarcity would become less important. Several reasons for this were shared, but the main idea was that we would in the future be able to collect more data, have better data management inside organizations, and better regulations would be in place.

Multiple respondents stated that data scarcity in general would not be an issue in the future, but data access could become this instead. Data is increasingly abundant in our current society. However, we are not always able to access those data. Noted was that data would become more of a currency between organizations which could be traded, with a centralized European body that would oversee this trade (ParticipantH3, 2024). However, one respondent also said that he was confident that there would be better data management laws and regulations, fixing this issue (ParticipantH2, 2024).

Other respondents stated that in the future organizations would have better data management, this would eliminate the reason of data scarcity for the application of TL (ParticipantG2, 2023). Additionally, advances with technology could also provide a shift towards the better availability of structured data for training purposes (ParticipantA1, 2023).

Lack of computational power

The second most dominant reason identified for the use of TL in organizations is the lack of computational power. Most respondents ranked this reason in the second or first position, except for all respondents employed in the healthcare industry. In average the lack of computational power was ranked as fourth reason in this industry, as seen in tables B.3, B.4 and B.5.

The reason for this is that this sector gives priority to different reasons for the use of TL. One respondent even considered the lack of computational power as a non-issue, he stated that whether you are applying traditional ML methods or TL, you would need the same amount of computer power.

> *"Lack of computation power is, well, almost a non-issue, I would say, because when you talk about computers, you're talking about floating point operations and VRAM that you use, those kinds of specs, and that remains the same. It doesn't matter (ParticipantH2, 2024)."*

Lack of computational power might pursue this industry in some cases, but there are other reasons that lead to the implementation of TL more often. One respondent stated for example that a consensus already exists between hospitals that they will not invest in their own server capacity. When they are designing AI applications, they keep this in mind from the start.

> *"The reason why I haven't rated lack of computational power very highly now is because it's essentially already a consensus that it's not a very good idea for healthcare institutions to start calculating things themselves (ParticipantH3, 2024)."*

Other industries considered the lack of computational power as more important. Respondents from the finance industry both mentioned that this could be a reason to implement TL instead of traditional ML (ParticipantF1, 2024), (ParticipantF2, 2024). One respondent also mentioned that the need for more computational power to train and run AI applications is also a expense issue. So, to keep expenses within budget, they were more likely to choose a TL solution (ParticipantF1, 2024). Another respondent stated that they used TL mainly for the application of LLM's. Lack of computational power was their reason of why they did not develop an LLM themselves (ParticipantF2, 2024).

> *"This constraint is a common challenge faced by organizations, and it not only poses difficulties but also adds to the expenses associated with implementing transfer learning (ParticipantF1, 2024)."*

> *"The current reasons why we use pre-developed LLMs are indeed lack of computational power. I think that's the main one (ParticipantF2, 2024)."*

The expense issue for lack of computational power was also confirmed by another respondent, employed in a general industry. He stated that since their organization's core business was not the development of ML models, they preferred the use of pre-trained foundation models (ParticipantG4, 2024). This saved the organizations some resources.

> *"Lack of computational power, combined with the fact that we simply don't want to allocate resources to it. Again, because we believe that every euro spent should ideally go directly to our core business (ParticipantG4, 2024)."*

Another reason that was given for lack of computational power as a driving force for the application of TL was the growing size of ML models. For example, when a respondent was previously building models himself, he could do everything on-premise. But with the introduction of gigantic foundation models, he expected that to become more difficult in the future (ParticipantG2, 2023).

Just like with data scarcity, respondents also predicted the lack of computational power to become a less important drive for TL in the future. Multiple reasons for this have been identified. One respondent stated that he expected the costs of GPU's to go down and performances to go up. This would make a lack of computational power less relevant in the future (ParticipantG1, 2023). Another respondent noted that she simply saw a lack of computational power as a hurdle that we would overcome in the future (ParticipantH1, 2024).

Maintaining privacy

Two reasons have the same overall current ranking of '3.6', as seen in table 4.2. One of those reasons is the application of TL for maintaining privacy. Respondents were again divided on how to rank this reason. Respondents from the healthcare industry all ranked maintaining privacy as reason number 2, as seen in table B.3, B.4 & B.5. Whereas all other respondents rated it lower.

Respondents noted that privacy is considered extremely important in the healthcare industry. Especially when working with patient related data. Data is then often anonymized or pseudonymized. TL could solve the problem of not being allowed to share data among hospitals, for instance by applying federated TL solutions (ParticipantH1, 2024), (ParticipantH2, 2024). However, one respondent also stated that the question of how this should be legally executed is still actively being addressed. She did not know whether a definitive solution was yet found (ParticipantH1, 2024).

> *"The advantage of transfer learning is that hospitals can share models without direct data transfer. This bridges privacy and legal limitations (ParticipantH1, 2024)."*

Another respondent confirmed that for the healthcare industry, privacy is a major concern of which Federated TL could be of help.

> *"When collaborating with various hospitals, for example, it is important to use privacy-preserving methods, such as federated learning (ParticipantF2, 2024)."*

A different respondent stated that even though federated TL solutions could theoretically address privacy issues, it would be difficult to get approval from privacy officers to share models with other institutions. A reason for this was that it is sometimes still possible to recover training data from a ML model (ParticipantH2, 2024). Furthermore, the idea of federated TL solutions for healthcare has existed for quite a while, but has not come out of the research phase yet, as stated by a respondent (ParticipantG3, 2023). This respondent also said that he however expected federated TL to become a viable practical solution in the future. Another respondent did not share the same positive look on the future and went as far to say that the healthcare industry was a horrible sector, in which good ideas such as federated TL framework were always limited to small initiatives that would never receive enough attention (ParticipantF2, 2024).

In the finance industry, maintaining privacy as a reason for the use of TL is ranked lower, as seen in tables B.6 and B.7. This does not mean that the finance industry considers privacy of individuals and organizations as unimportant, but it shows that maintaining privacy is currently not a major incentive for applying TL. One respondent stated that privacy is important for the financial industry and that sharing sensitive personal information is heavily regulated. For instance, for the development of ML models, certain agreements on how to share data exist.

> *"However, there are instances where certain agreements allow financial services to share data with third-party companies, especially for purposes like anomaly detection and combating money laundering (ParticipantF1, 2024)."*

Another respondent noted that privacy was not a major concern for them because they only use internal data, also for TL purposes. They currently do not share data or models with third-parties. Everything follows their internal privacy guidelines, so having maintaining privacy as a reason for implementing TL is not relevant for them (ParticipantF2, 2024).

Regarding the application of federated TL solutions in the finance industry, one respondent stated that it is a potentially viable solution which holds promise, especially regarding financial crime and money laundering. However, he did not know of any current implementations. Federated TL was probably still in the research phase. Another respondent agreed with this opinion, but repeated that they did not have any use for federated TL since all their data and operations are held internally (ParticipantF2, 2024).

> *"Federated learning emerges as a potential solution in this context, and while it hasn't been widely implemented, it holds promise, particularly in addressing the specific challenges of financial crime and money laundering. There have been initiatives, such as a project from the University of Delft, exploring the applicability of Federated Learning in this domain (ParticipantF1, 2024)."*

In addition, a respondent stated that federated TL in the finance industry would be difficult to implement because of regulatory issues. While the solution might work well for fighting financial crime and fraud, it would require intense collaboration between financial institutions. The respondent raised questions about how a regulatory framework for this would look like. He expected it to be extremely difficult (ParticipantF1, 2024).

Other respondents from different industries stated among other things that TL was not employed for privacy reasons. One reason for this was that all used data was not personal, pictures of fruit did not have the need for privacy (ParticipantG1, 2023). Another respondent stated that he did know what federated TL was, but that they also did not have a need for it since they only worked with internal data (ParticipantG4, 2024).

When asked about how this reason would change in the future, there was a general consensus that maintaining privacy would become more important. A reasons given for this was for example the expansion of privacy related laws and regulations such as the GDPR (ParticipantG3, 2023). Privacy in general would become more important, but it would also become a bigger drive for the use of TL.

One respondent however stated that privacy in the future would become way less important. Currently, people still regard data privacy as an important right, but in the future he expected this to become less of a concern. People would begin to understand why and how their data is being used, and be more okay with it. He considered privacy as a luxury.

> *"I actually think it's a sort of expression of extreme luxury and prosperity to be able to say my data, it's not used for anything. I think you really have to make a lot of effort to achieve that, and I don't think it's a problem at all. So that's why you might think that privacy is a very current topic now, but in the future, it hardly plays a significant role (ParticipantH3, 2024)."*

Lesser environmental impact

Lesser environmental impact as a drive for the use of TL was regarded just as important as maintaining privacy, as seen in table 4.1. Currently however, this was not a main reason for the use and development of TL applications for organizations. This reason was often discussed together with TL as a reason for the lack of computational power. The less computer power someone needs or uses, the lower the environmental impact will be.

One respondent noted that when they would develop AI applications, they would always keep the environmental impact in mind. However, they had not yet looked at TL as a way of lowering emissions. They mostly looked at it from a cost perspective. When you use less computational power for your applications, your projects can be completed with a smaller budget. At the same time, this also keeps emissions low. He stated that they would run most of their AI applications in Belgium, because the carbon footprint was calculated to be the lowest in that country (ParticipantG4, 2024).

Two respondents stated that in their industry environmental impact was not a driving force for the use of TL. However, in their personal opinion, it should be. One respondent stated that with the rise of bigger ML models, more attention should go towards environmentally friendly AI development. TL could in his opinion be a solution for this. (ParticipantG3, 2023). This opinion was shared by another respondent who stated that environmentally friendly AI development was important, but hospitals did not care much about that yet.

> *"For example, I personally find the environment very important, but I notice that hospitals are not yet very focused on it (ParticipantH1, 2024)."*

Environmental impact for AI development was also noted as being hard to measure. Questions arose about how someone could measure the emissions made by for instance a foundation model. Because TL applications are often cloud-based and running on external hardware, it is difficult to determine the environmental impact. Because it is so hard to measure, a respondent noted this reason as a lesser driving force for the use of TL (ParticipantA1, 2023).

The last reason given why lesser environmental impact was not a driving force for TL was that in the automotive industry, the main way of becoming more environmentally friendly was mostly focused on improving the car production processes and engines. Impact on the environment of AI was regarded as very small in relation to other emissions in this industry. Therefore when they focused on the climate, they did not give much attention to AI and TL.

> *"In comparison, the environmental impact of the AI activities within our sector is likely minimal, and as a result, it doesn't receive as much attention (ParticipantA1, 2023)."*

All respondents agreed that in the future, the use of TL for a lesser environmental impact would probably become more important. Since organizations are increasingly looking to reduce their carbon footprint, reducing their AI related emissions would be a logical step to take. One respondent stated that they have to adhere with an increasing amount of environmental guidelines, based on this he said that lesser environmental impact would definitely become more important.

> *"The growing emphasis on sustainability in business practices, including the environmental impact of operations, is becoming increasingly important*
> *(ParticipantF1, 2024)."*

> *"Environmental impact is becoming increasingly important. Because we have to comply with certain environmental guidelines, so to speak, as a bank. So this is definitely a point. (ParticipantF2, 2024)."*

Personalized AI

The last reason mentioned earlier, which could encourage the use of TL in organizations, is the ability to create more personalized AI applications. On average, this reason was ranked the lowest of the five initial reason, with a score of '4,1', as seen in table 4.1 Most respondents found the idea of more personalized AI an interesting concept, but it was not clear for them yet how they would apply this in their industry.

One respondent employed in the healthcare industry however saw great opportunities in TL for the creation of personalized AI. At the current moment, he did not know of any applications, but he stated that personalized AI applications, made with TL, could potentially make surgeries more safe (ParticipantH3, 2024). General pre-trained models could be personalized with data of a patient who is about to undergo a complex procedure. A model would then be fine-tuned on all kinds of patient data such as lab results, blood types, medications and previous medical history. By making use of online TL, feeding the model with real-time data from the operation table, the surgeon could receive instant feedback while performing surgery. This was still a theoretical concept, but active research on this was being performed. In the future, he was convinced that creating personalized AI would become an important incentive for the use of TL.

> *"So, I find that last part, the personalized part, really interesting because a generic model is of course very nice, but ultimately you always have to apply it to individual patients, and there I notice quite a big difference in the performance of some models (ParticipantH3, 2024)."*

A different respondent who conducted his PhD on personalized AI and TL in the healthcare industry, stated that in his research, personalized AI based on TL was a favourable approach resulting in efficient models. Unfortunately however, he did not know of any practical implementations of this (ParticipantF2, 2024). Personalized AI based on transfer learning has not progressed beyond the research phase.

> *"The ideas are good, the expectations were also high, and I think very little has come of it in practice in terms of transfer learning because in practice, it is hardly used at all (ParticipantF2, 2024)."*

Other respondents from the healthcare industry stated that they did not have any experience with TL for personalized AI (ParticipantH1, 2024), (ParticipantH2, 2024). They could see some potential in it when the concept was explained, but admitted that they lacked knowledge about this subject to say something meaningful.

In the finance industry, personalized AI was by one respondent identified as a current incentive for the use of TL, in particular regarding the fine-tuning of foundation models. The respondent stated that they used TL to personalize their LLM's (ParticipantF2, 2024). Another respondent stated that TL to enable personalized AI was currently not an important concept in this industry.

> *"I perceive the concept of personalized AI as a secondary consideration, ranking it at number five (ParticipantF1, 2024)."*

In the long term, this reason for TL was perceived to become less relevant by one respondent only. He noted that foundation models would become proficient enough in personalization without the need for fine-tuning. He was of the opinion that TL might become redundant because of this.

> *"...and personalized AI, I think that in the long term, it becomes less important because those models become very good at personalizing anyway, as you can already see with LLM's, they are really excellent at it (ParticipantF2, 2024)."*

Another respondent stated that personalized AI made with TL was an interesting concept, but that he saw significant privacy related hurdles before this could be implemented in organizations. He was of the opinion that his organization could not legally build these personalized models.

> *"I would find that extremely interesting, but I think with the current AI legislation and data privacy, it just becomes very difficult to achieve that (ParticipantG4, 2024)."*

Finally, someone noted that he expected organizations to apply personalized AI models in the future, based on fine-tuned foundation models. An example given was a personal assistant trained on personal data of employees (ParticipantG2, 2023). Overall, except for one respondent, everyone agreed that the possibility to create personalized AI would become a bigger incentive for the use TL in organizations in the future.

Newly identified reasons for TL

In addition to the five reasons for the use of TL previously identified in the theoretical framework of this thesis, several new reasons have been identified based on the interviews with professionals.

The first new reason why organizations would apply TL in their operations is for overall efficiency of the development and execution of AI. Multiple respondents stated that this was an important incentive for them to choose for a TL approach (ParticipantG2, 2023), (ParticipantG3, 2023) & (ParticipantG4, 2024). This reason combines multiple previously stated reasons, but is still important to note. Overall efficiency can be a major incentive for why an organization would execute a particular strategy. Respondents stated that AI applications could be developed significantly faster when general foundation models can be fine-tuned, instead of developing and training them from scratch.

> *"And our entire frameworks are actually built in such a way that it's easy to quickly move from a concept or an idea to a workable model in production (ParticipantG4, 2024)."*

Another example related to efficiency was cost-efficiency. A respondent stated that he had the ability to create a smaller model from scratch that was just as accurate as a fine-tuned foundation model. But that it would cost more time and money to do that. For this reason, they would prefer a TL approach.

> *"Firstly, developing software around an existing transfer learning model is much simpler than building from scratch and fully training new models. This means it takes less time and the direct costs are lower (ParticipantG1, 2023)."*

The second newly identified reason for the application of TL in organizations is the scarcity of relevant skilled personnel. Multiple respondents stated that they considered this a strong reason for TL use. While it would be desirable for organizations to develop all their models from scratch, not having enough relevant people to do this pushes them to implement a TL approach (ParticipantG3, 2023). It was considered as too big of an investment to make. One respondent from the automotive industry confirmed this.

> *"Finding individuals with the right skills to train large models can be challenging. Simplifying the process to accommodate the available talent and making it more time-efficient becomes crucial (ParticipantA1, 2023)."*

By using TL, organizations can develop and deploy AI applications in an easier way. There is less need for in-depth knowledge about ML development, since fine-tuning is considered easier than developing a LLM yourself (ParticipantG1, 2023).

Finally, one respondent stated that the reason TL was used often in their AI approaches was because of a bottom-up approach. Employees would develop simple AI solutions in their spare time. Because a lot of AI solvable problems are not unique, TL is often used. When you encounter a problem, there is a high chance that someone has already thought of a solution and made it available. A lot of organizations face the same challenges and there is no need to reinvent the wheel every time (ParticipantA2, 2024).

4.2.4 Challenges regarding TL

Multiple challenges regarding the implementation of TL in organizations have been identified. They are explained in this section, together with recommendations given by respondents on how to overcome them.

Laws and regulations

The first challenge that was brought up by multiple respondents was concern about how to adhere to laws and regulations. In particular regarding licensing and data accessibility. Respondents from multiple industries noted that they often had to ask themselves two questions. Firstly, is the organization allowed to use an external model? And secondly, is the organization allowed to access and use the needed data to apply TL? Laws and regulations regarding this have become stricter in the last couple of years, limiting organisations in the collection of data and their freedom on choosing external models.

Often, rules regarding how an organization is allowed to use an external model are perceived to be unclear. A respondent stated that they had issues in determining how they were allowed to use a model (ParticipantA1, 2023). This could lead to a situation in which a model that suited their needs would not be used, simply because they could not determine if it was allowed. Another respondent from the automotive sector confirmed this and stated that it was extremely important for organizations to check their licensing (ParticipantA2, 2024).

> *"Sometimes things are unclear about licenses, mostly on the industry side. So, when you have a pre-trained network, what kind of license is it? How can I the model apply it and for what? Am I allowed to apply it? (ParticipantA1, 2023)."*

Another respondent confirmed this statement and added that developers of pre-trained models often do not think about how they license their products. When for example a model from a research paper is published online, it is unclear if the weights and the model can be used for commercial purposes (ParticipantG3, 2023).

> *"But when you talk about transfer learning, I actually want those weights. You want the knowledge from the model. That's often where the problem lies. Because people who publish those models often don't think about that. And then you run into the issue that those weights are indeed on*

the internet. But it's not entirely clear whether you're allowed to use them or not (ParticipantG3, 2023)."

Respondents from the healthcare industry stated that they could also confirm this challenge (ParticipantH1, 2024), (ParticipantH3, 2024). Every model that is used in the healthcare industry needs to be properly licensed when it is shared or used. For example, developers need to explain why a model would give certain results. If a model gives incorrect results and a patient dies, hospitals and developers will probably be sued (ParticipantH3, 2024). This shows that licensing is crucial for safety and ethics, but it can limit innovation (ParticipantH1, 2024).

> *"But when you look at healthcare, you often need to be able to justify yourself. Why a model gives the results it gives. And if you then say, "Well guys, we didn't detect cancer properly. But yeah, we just grabbed that model from somewhere." Then you'll end up in a lawsuit.(ParticipantH3, 2024)."*

Regarding finance, a highly regulated industry, a respondent noted that developing TL models was often not worth the amount of work because of strict regulations. Models they developed underwent validation by a distinct and independent department before they could be deployed into production. Getting fine-tuned external models through this validation process was not worth the work (ParticipantF1, 2024).

Risk of bias

If an organization does have the right licenses to use an external model, some new challenges can be identified. One of which is the risk of introducing bias into your models. In general, fighting bias in AI is perceived as an important thing to do. However, when developing AI applications based on TL, specifically focused on the fine-tuning of foundation models. Things can get a little more complicated. Since developers are using a pre-trained model which they have not build themselves, they do not know which kind of biases may exist in this model.

Multiple respondents stated that the risk of introducing bias in AI models with the use of TL is a challenge (ParticipantG1, 2023), (ParticipantG2, 2023), (ParticipantG3, 2023) & (ParticipantF1, 2024). It is often unclear what the training data consists of, and whether perceived biases are intentional or not. It is important for organizations to be aware of the possibility of these biases, and to implement strategies to mitigate them.

> *"Since these models learn from extensive datasets, including sources like Wikipedia, they are susceptible to inheriting biases prevalent in the training data, particularly regarding gender, race, and other factors. Managing and mitigating such biases become critical considerations in the development and validation of these language models (ParticipantF1, 2024)."*

Respondents stated that it is difficult to avoid bias in general, because foundation models are trained on such vast amounts of data. Making sure that the training data is a fair distribution of the real world is difficult, but important (ParticipantG1, 2023). Nonetheless, this makes it almost impossible for the developers of these models to identify every single bias (ParticipantF1, 2024), (ParticipantG1, 2023), (ParticipantG4, 2024).

> *"The amount of data that goes into these models is so vast that it's impossible for the developers to verify if all the data is correct and doesn't contain any unintended biases (ParticipantG4, 2024)."*

This makes it very important for organizations to make sure that they implement strategies to mitigate the risk of bias in their TL based AI applications. Especially because it is the responsibility of the party that is using, applying or selling the TL application to make sure that a model is safe. You need to test your own fine-tuned model, since it is eventually going to be your property (ParticipantG3, 2023).

Multiple respondents stated that they implemented measures to combat bias. A respondent from the finance industry stated that their organization already had developed a testing framework for TL based AI applications, mostly focused on text generation. They build a pipeline in which their internal benchmarks were used to test fine-tuned models (ParticipantF2, 2024). Another respondent said that they only used external models which were tested based on external benchmarks (ParticipantG4, 2024).

> *"For the LLM's part, generative AI, it's naturally more difficult because it's different. We have now taken the first steps there. It's just a pipeline with which we have our own internal benchmarks. Where we pay attention to certain things that could influence responsible AI (ParticipantF2, 2024)."*

> *"No, it does mean that in the preselection of larger models, we do look more closely at models that have a low bias on certain benchmarks than others (ParticipantG4, 2024)."*

Technical challenges

Other challenges respondents spoke of can be classified under technical challenges. These are all obstacles that could come to light during the development, implementation, or maintenance of TL based AI applications.

The first challenge a respondent noted was related to operational deployment of TL solutions. Organizations should consider whether a model needs to operate in real-time or can function in a batch processing mode. The choice of this impacts the overall solution of your AI application and architecture. He noted that if a model should for instance process data in real-time, smaller sized AI applications are often preferred because of quicker inference time (ParticipantF1, 2024). Using fine-tuned foundation models is then often not a preferable solution.

Another respondent noted that interoperability can be a significant challenge for the implementation of TL, especially for the healthcare industry. On one hand this industry is increasingly working on standardizing their software and practices, but on the other hand they are still actively using the fax machine and hand-written notes. Additionally, a very broad spectrum of technological standards are still accepted, which toughens the broader implementation of TL technologies and AI in general (ParticipantH3, 2024). For instance in the development of an industry wide database of TL models. If every healthcare organizations uses significantly different data types and tools, collaboration and model sharing is hard.

> *"That we can at least agree on which standard data formats we use. Perhaps a sort of expressed preference for certain analysis tools or programming languages, for example (ParticipantH3, 2024)."*

The third challenge identified is related to a shortage of computational power. Even though computational power was also identified as a reason for why organizations implement TL based AI applications, a shortage of computational power can also hinder better TL implementation. A respondent stated that this challenge was specifically related to the fine-tuning of large foundation models, since fine-tuning these models still requires high GPU capabilities and storage capacity.

> *"Computational power remains a significant hurdle, particularly when dealing with larger and more sophisticated LMs. While models like BERT might be trainable on a standard laptop, the new generation of LMs demands substantial resources, both in terms of GPU capabilities and storage capacity. This poses a challenge for fine-tuning or further training with proprietary data (ParticipantF1, 2024)."*

The last technical challenge that has been identified is the challenge that fine-tuning pre-trained models is not always as easy as it looks. For example, a respondent noted that they sometimes dealt with catastrophic forgetting of models. During the fine-tuning process, a model would sometimes forget a lot of knowledge, without it being clear why that was happening (ParticipantG4, 2024). This is something that organizations should keep in mind.

Organizational mindset

Overall, there was one reoccurring theme which can also be classified as a challenge regarding the implementation of TL in organizations. This is not a technical or directly AI related challenge. However, it is a challenge of the general mindset of organizations that want to use and implement AI and TL. Multiple respondents from different industries stated that the mindset of people and organizations was the biggest challenge. A respondent stated that the end users of TL products and AI in general often do not understand how a product works. This makes it difficult to give them an incentive to actually use TL applications in their day to day operations (ParticipantF2, 2024).

For instance, a respondent from the healthcare industry stated that there is often discussion in hospital about when an AI model is good enough. Or the doctors that are the end user of a TL application do not actually want to work with the tools because they lack the enthusiasm (ParticipantH1, 2024). This can significantly slow down the implementation of AI and TL in organizations.

Another respondent, also from the healthcare industry, noted that healthcare related organizations often perceive themselves as unique and better than others. He stated that because of this, it was difficult to organize collaborations between parties. Healthcare organizations should realize that while their culture may be unique, their data and tools often are not. If they open themselves up for more collaboration, TL tools can be better developed and implemented more easily industry wide.

> *"Yes, the biggest challenge is indeed the tendency of every hospital, every healthcare institution, to see itself as unique. That's a bit, well, that's quite laughable, of course. Every hospital I speak to, especially the academic medical centers, they are convinced that they are truly unique and fantastic and brilliant. And that all others really shouldn't even be allowed to come close to their level. That's a kind of mindset that we actually need to move away from (ParticipantH3, 2024)."*

This was also confirmed by another respondent, he added that if TL was to be implemented in a whole industry, organizing things top-down can be a recipe for disaster. It was for instance easier to develop and integrate TL applications in the finance industry because the business model allows for this. Healthcare organizations have strict budgets and TL based AI applications are always confined to either very small initiatives that fail because they are not given enough attention and funding, or initiatives that are too big because the industry immediately tries to organize it sector wide (ParticipantF2, 2024).

Additionally, another respondent also stated that the lack of a business case can be a reason for why TL or AI applications are not implemented. She described an example of an Intensive Care Unit (ICU) that did not want to integrate an AI model which could predict when a patient could leave the department. The model could significantly reduce the amount of time in which a patient was in the ICU. This however meant that the head of the ICU would receive less earnings because his occupation was lower (ParticipantH1, 2024). The example shows that even though an AI application can overall improve an industry, it can be haltered because of different interests.

Finally, one respondent stated that he did not know of any challenges related directly to TL as a technique. He did not perceive organizational challenges to be TL specific, but applicable to every type of new innovation. Organization could for instance not implement a good solution because of time constraints, a lack of interest, or different priorities (ParticipantA2, 2024).

4.2.5 Gartner Hype Cycle positioning

In the final part of the interviews, respondents were asked on which location they would place TL on the Gartner Hype Cycle. The given answers are visually represented in figure 4.1. A clear divide in the answers can be observed, the majority of respondents placed TL either just before the 'Peak of Inflated Expectations' or on the 'Slope of Enlightenment'. Answers have been placed inline, but all represent more or less the same location on the cycle. Furthermore, one respondent placed TL in a different phase of the Hype Cycle, just past the 'Peak of Inflated Expectations on the slope to the 'Trough of Disillusionment'. He expected that a small number of professionals already know the potentials of TL, but most do not, which will first lead to a number of unpractical applications, before real productivity could be reached.

Please note that the total amount of recorded answers on figure 4.1 exceeds the number of conducted interviews. This is the case because some respondents gave a double answer, one from an industry-wide view on TL and one from a technical point of view.

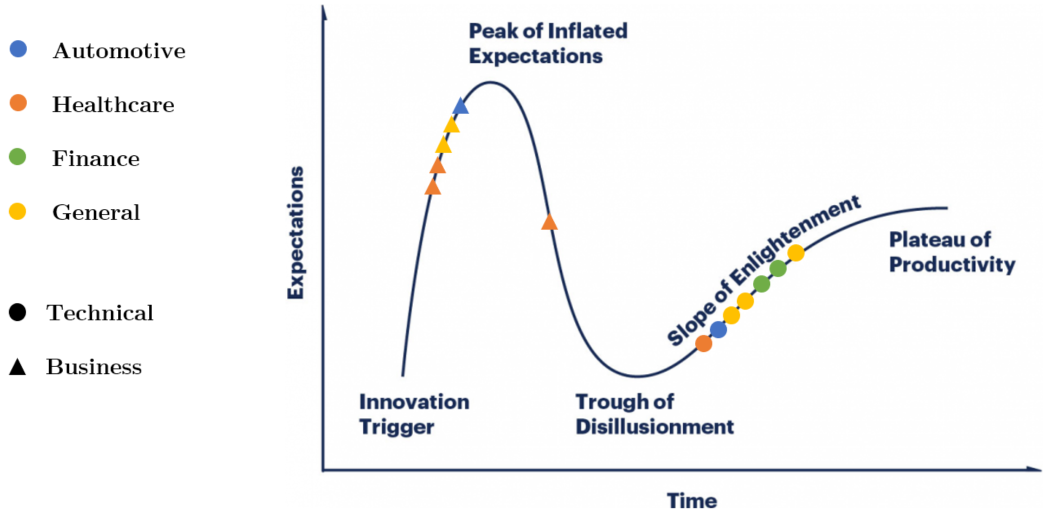


Figure 4.1: Positioning of TL on the Gartner Hype Cycle.

Industry-wide views generally placed TL right before the peak of the cycle. TL was perceived as a useful AI strategy which is gaining attention, but has not yet reached it's peak of expectations. Once people have a better understanding of what it can do, it could garner a significant hype. However, not yet fully identified setbacks such as data and other technical limitations will likely follow, deflating the expectations (ParticipantA1, 2023). One respondent also noted that TL could potentially lift on the back of the current Generative AI-hype, which in turn can inflate the TL expectations of an industry (ParticipantG2, 2023).

Another group of respondents placed TL in a much different phase on the Hype Cycle. They perceived TL to be on the 'Slope of Enlightenment'. These respondents were looking at TL from a more technical point of view. Respondents stated that there exists a growing awareness of the practical applications of TL, with people becoming increasingly familiar with how to employ it (ParticipantF1, 2024). Together with the rise of fine-tuning of foundation models, TL is finally starting to develop more useful applications for organizations (ParticipantF2, 2024). However, TL was not perceived as a new technology.

The different views were explained as follows: more technically educated people who have hands-on experience with TL are already aware of the possibilities and limitations of the technology. Because of this, they would place TL more on the 'Slope of Enlightenment'. The situation was perceived different for professionals who are less technologically educated and are more focused on the business side of organizations and the industry as a whole. They were perceived to only recently have learned of TL and it's possibilities. For them, the technology is new and exciting, potentially inflating the expectations.

Additionally, multiple respondents noted that they found it difficult to place TL on a location on the Gartner Hype Cycle since they did not perceive TL to be a hype or follow a hype trend (ParticipantG1, 2023), (ParticipantG3, 2023). TL was perceived only as a means of reaching your AI goals, not a driver for applications itself. Developers know what it can do and will use it accordingly, but it would not have a lot of media attention.

> *"But I also don't expect there to be such a valley compared to the peak. I think rather that for transfer learning specifically, that's something that has proven itself. There's not a huge amount of hype around it (ParticipantG1, 2023)."*

> *"If I had to place it somewhere, then it's still before the hype indeed. I've never really noticed that it's a hype (ParticipantG3, 2023)."*

Another respondent did think that TL followed the Hype Cycle and believed the 'Peak of Inflated Expectations' was reached around 2018. He stated that the expectations for TL at that time were different and much higher (ParticipantA2, 2024).

> *"I think so. That was, I mean, that was my feeling. When there was a new problem, everybody said, ah, okay, we just take a network trained on ImageNet and then we use it on audio data. No problem (ParticipantA2, 2024)."*

Finally, one respondent placed TL just over the 'Peak of Inflated Expectations' on it's way down into the 'Through of Disillusionment". He expected people who engage with TL in depth to have heard of TL, and to know the possibilities and limitations. However, in general, many people who are less familiar with the subject can be confused about the capabilities of TL (ParticipantH3, 2024).

Chapter 5

Discussion

This chapter includes a summary of the results together with the evaluation and linkage to the theoretical framework. Additionally, implications and limitations of the research are discussed, together with practical recommendations for organizations. The research questions of this thesis was:

How is Transfer Learning currently being adopted across different industries and what are challenges and opportunities based on the identified TL approaches?

To assist the research question, four hypotheses have been formulated. The interpretation of the results is organized around them.

5.1 H1: Varying adoption of TL

The first hypothesis was: The adoption of TL varies across different industries, with some industries being more receptive than others.

In total, 11 participants have been interviewed employed in 7 different industries. 4 of which were working outside of the three main focus industries. TL has been identified to be applicable in all sectors. Every respondent was able to give examples of how they applied TL in their respective industry. TL appears to be abundantly used for more common tasks, often referred to as 'low-hanging fruit.' Examples of this are applications for computer vision and NLP. The less specific an AI use-case is, the higher the chance that it has been done before, which in turn makes it easier to apply TL, instead of developing solutions from scratch.

Based on the observations, it seems that the current state of adoption varies per industry. Some industries are more susceptible to the adoption of TL than others. Therefore, the hypothesis that the adoption of TL varies across different industries, with some industries being more receptive than others, is accepted.

Automotive

The automotive industry shows a great susceptibility to the implementation of TL for a wide range of use-cases. TL can be used for, among other things, autonomous driving applications and smart manufacturing. This industry currently implements

TL in a reactive way, driven by the needs of development teams. No broader AI strategies tailored to TL have been identified, however, the automotive industry does seem open to it. In the future, this industry is likely to see an even greater integration of TL into its operations. With the ongoing evolution of autonomous driving systems, smart manufacturing processes, and connected vehicle technologies, the demand for TL based AI solutions will only increase. While the current application of TL is mostly driven by specific project needs and focused on 'low-hanging-fruit', a broader strategic vision on TL can lead to more efficient AI development and a greater scalability of AI applications. Furthermore, the automotive industry is currently often uncertain about licensing and regulations regarding TL. A clearer understanding and establishment of regulatory frameworks specific to TL and AI-model reusage would provide much-needed guidance and assurance for automotive organizations looking to implement TL-based AI solutions.

Finance

Likewise, the finance industry also seems receptive to the implementation of TL. Even though not many actual TL based applications have been identified, they have also not been proven impossible. The reason why financial organizations currently choose different AI methods are because they often possess the data and resources to develop models from scratch. Current AI applications which do use TL are mostly focused on the the implementation of NLP in financial services. With the newly available LLM's, it is expected that the use of TL will see a big increase in the upcoming years. The finance industry is currently developing, testing and identifying new applications for this. These are for instance focused on customer relationship. It is expected that almost every custom-trained NLP application will be based on TL, as the cost-effective training of LLM's is not a feasible solution. TL can also make the development of risk management and fraud detection algorithms easier, for instance when an organization is looking for international expansion and needs be able to handle different country-specific regulations. An AI strategy which is more focused on the reusing of previously developed models can make the development process more efficient and cost-effective for financial organizations.

Healthcare

Regarding the healthcare industry, it has been identified to be the least receptive to the adoption of TL in practice. Certain TL applications have been identified, but these are mostly confined to research settings. Despite having numerous ideas and plans for how TL could improve healthcare, regulations and resistance to change often seem to hinder actual adoption. Additionally, certain levels of ignorance among those who need to use TL in practice further contribute to this aversion. This is a problem, because TL can in theory make a big impact on making the healthcare industry more ready for the future. TL for computer vision has the ability to detect more rare diseases and improve diagnostic accuracy, while TL for NLP can enhance medical documentation, streamline administrative tasks, and facilitate more efficient communication between healthcare professionals and patients. Failure to successfully embrace TL may hinder the healthcare industry's ability to address pressing challenges such as rising healthcare costs, increasing patient volumes, the growing complexity of medical data, and the decreasing number of available healthcare personnel. TL can make a difference. However, since the barriers to adoption are primarily organizational rather than technical in nature, similar to the two other industries, AI strategies tailored to TL are critical for this industry.

If organizations want to expand their use of TL in practice. They have to overcome some industry specific and general challenges, which will be explained in the following hypothesis. By taking away these impediments, they clear the road for a broader implementation of TL in their AI strategies.

5.2 H2: Influential factors on TL implementation

The second hypothesis was: Factors such as data availability, privacy, industry-specific needs, industry-specific regulations, and technological infrastructure play a significant role in the extent of TL implementation across different industries.

5.2.1 Why organizations use TL

Seven specific reasons for the application of TL in organizations have been identified. These reasons all play a role in the extent of TL implementation across different industries and organizations. The identified reasons are:

1. Data scarcity
2. Lack of computational power
3. Maintaining privacy
4. Lesser environmental impact
5. Personalized AI
6. (Cost) Efficiency
7. Lack of skilled personnel

Some of these reasons are more decisive than others, as seen in table 4.2. Of the five reasons that have in advanced been stated in the theoretical framework, all have been confirmed by the respondents. 'data scarcity' and 'lack of computational power' are currently considered the most important for why organizations apply TL. However, in the future these are expected to become less important. It is expected that a focus on better data-management and the improvement of computing power will minimize these challenges. On the contrary, 'Maintaining privacy', 'lesser environmental impact', and 'personalized AI' are expected to become a more important reason for organization to apply TL. Privacy is an increasingly important subject in our civilization, when TL proves to help us create better AI applications without compromising personal data, it can become a more important reason. The environmental impact of AI is also a subject that is gaining more interest. Lastly, TL facilitates the creation of more personalized AI applications, this will become more important.

Additionally, two new reasons have been identified during the interviews. These are the use of TL because of (cost) efficiency and lack of skilled personnel. Organizations use TL because it can be cheaper and they need less developers to implement something. Since they have been identified later, during the interview process, they have not received a ranking.

Automotive

The automotive industry seems to mostly implement TL because of a need for more personalized AI models and data scarcity. Maintaining privacy and a lesser environmental impact are not big incentives for TL. This is because the automotive industry does not use a lot of sensitive personal data for their machine learning purposes. For instance, racing or smart manufacturing applications primarily rely on performance metrics and production data rather than sensitive personal information. Additionally, the automotive industry can make a bigger positive impact on the environment by actually reducing emissions from both cars and production processes, rather than focusing on the environmental impact of their AI. Therefore, the incentive for TL in lies more in the need for personalized AI models tailored to specific vehicle functionalities and operational requirements, as well as overcoming data scarcity challenges.

Finance

The biggest incentives for TL for the finance industry are data scarcity and a lack of computational power. Currently, TL is mostly used in this industry because of efficiency. It allows for faster training times, lesser personnel, and lesser data requirements. Most organizations in this industry are profit-driven, which means that a lot of decisions are made based on a cost-benefit analysis. If TL offers more gains for a lower cost, it is the favourable approach. In the future, some reasons are however expected to become more important for the finance industry. Because of stricter data regulations, new AI laws, and a bigger focus on the environment, this industry is expecting an increase in the use of TL for privacy and a lesser environmental impact. It is recommended for financial organizations to anticipate and adapt to these evolving trends by integrating transfer learning into their operations, not only for its efficiency benefits but also to address these emerging topics.

Healthcare

Unlike the other industries, in the healthcare industry, privacy is considered the second most important reason, after data scarcity. The healthcare industry is an industry where the proper handling of personal data is of great importance. This translates to how their reasons for TL are ranked. Even tough this industry currently does not have a major number of TL applications in production, incentives do exist. Therefore, while the healthcare sector may lag behind other industries in TL implementation, the potential benefits of enhanced privacy protection and improved data security serve as big motivations for future adoption and integration of transfer learning technologies. However, this industry has some major hurdles to overcome before they can fully reap the benefits of TL. In general, there is a need for greater openness to innovation within the sector.

In general, TL is often applied reactively, driven by identified use-cases rather than serving as a proactive catalyst for new application creation. In the innovation process, organizations will first identify challenges they like to address. To solve these challenges, they then explore various AI approaches, including TL. However, TL itself does almost never seem to be the starting point of ideation. Organizations do not start with TL in mind and then think of applications that could come from it.

5.2.2 Challenges related to TL in organizations

Additionally, four main challenges have been identified which can hinder successful implementation of TL in organizations. Respondents from all industries confirmed that these challenges were applicable to their respective industry. Overall, the challenges regarding TL implementation do not appear to be overly specific to TL itself. Laws and regulations are a challenge, but not only for the implementation of TL, the whole AI field is adapting to new AI regulations. The same can be said for the other reasons. Bias is always something to keep in mind during the development of AI, everything new that is to be implemented can be technically challenging, and the mindset of organizations can halter every new innovation.

This does not mean that these challenges are not important for the further expansion of TL in organisation. So, how can the industries overcome these challenges? It is important to note that this study has mostly focused on identifying challenges and opportunities, not on overcoming them. More research or a separate study focused on specific industries is recommended to delve deeper into the industries and create actionable plans to effectively address the challenges. However, some general ideas can be constructed from the results of this study as well:

1. **Laws and regulations**: The field of AI is rapidly changing, this leads to the creation of more laws and regulation. With AI laws becoming stricter, organizations have questions about how they can use TL. Licenses of external models are not always clear, the same applies to data access. If organizations lack clarity on permissible usage, they are less likely to engage in it.

   For example, the automotive industry has mostly expressed concerns about licensing of external models and laws and regulations in general. Currently, the task of checking if a TL application adhered to the rules is up to the development team itself. This can be a problem because engineers might lack the needed deep understanding of regulations to make well informed decisions. Implementing company frameworks and procedures for evaluating TL applications against relevant laws and regulations can mitigate this risk. These teams could consist of a wide range of legal experts, compliance officers, and engineers to ensure comprehensive oversight and compliance throughout the TL implementation process. This recommendation is applicable to the other industries as well.
2. **Bias**: Bias in AI is always a challenge. However, when with TL external models are used, an unclear data origin makes the limitation of bias in AI even more challenging. Organizations have a need for certified models and strategies on how to identify and eliminate biases when they develop new applications from external models.

   Overcoming all biases in TL is a difficult task, however organizations can and should implement several strategies to mitigate bias and promote fairness in TL based AI applications. The organizations from finance industry have for example adapted their internal benchmark pipelines for 'traditional' ML, to also fit external models. This would be an important recommendation for other industries as well.

3. **Technical challenges**: Organizations can face certain technical challenges during the development of TL based applications. Certain challenges such as catastrophic forgetting and interoperability can slow down development.

   Overcoming these challenges is definitely possible for all industries. One approach is to invest in research and development efforts aimed at addressing specific technical challenges associated with TL-based applications. All interviewed industries also noted that this was not the most important reason for why TL was not implemented. Most reasons for this were more organizational.

4. **Organizational mindset**: Organizations can be slow or unwilling to implement TL because they lack awareness, enthusiasm and are ignorant. When an industry has a closed mindset for the implementation of new technologies, it is extremely difficult to reach something.

   This challenge is applicable to every industry, but most prevalent in healthcare. The healthcare industry, by its nature, tends to be more risk-averse and cautious when adopting new technologies due to concerns surrounding patient safety, regulatory compliance, data privacy, and lack of awareness under medical professional. To address these challenges, it is essential to prioritize education and awareness initiatives to increase understanding and enthusiasm for TL among healthcare professionals and decision-makers. Not only ML developers should know about the potentials of TL, but also managers and medical personnel.

Based on the identified challenges and opportunities, it has been proven that there are numerous factors which influence the extent of implementation of TL in organizations and across different industries. Therefore, this hypothesis is accepted.

5.3 H3: Varying TL approaches

The third hypothesis was: Implemented approaches of TL vary across industries based on specific AI needs.

TL approaches of industries vary to a certain extent based on specific AI needs. Certain industries might face challenges and requirements which influence the choice of TL approaches. For instance, the healthcare industry is less likely to use unaudited external models for their TL applications in production environments in comparison to the automotive industry, since it places a higher priority on privacy and data security, whereas the automotive industry may prioritize efficiency. The financial industry also deals with tailored laws and regulations in its field, effecting their choice of TL approaches.

However, asking which specific categorizations as approach of TL organizations use most deemed to be an unimportant question. Since organizations look at TL as a tool to solve problems, they did not care too much how they solved them, as long as it was successful. Whether an inductive or transductive approach to TL was taken, did not matter for a whole industry, but was based on the specific requirements and constraints of the use case. Analogous to a carpenter's tools, different TL approaches serve different purposes. Just as a carpenter selects tools based on the task at hand, organizations choose TL approaches based which can be best used to solve their use-case. Therefore, this hypothesis is partially accepted. TL approaches vary depending on specific AI needs, rather than being strictly defined by industry; they are more closely tied to individual use-cases.

Nonetheless, in general, one approach of TL deserves a special focus: the fine-tuning of foundation models. Organizations from all industries recognize the great potential of the newly available LLM's and their ability to be fine-tuned for specialized applications. This really is a new application that has never been available before. It is expected that this will greatly expand the application of TL in all industries. Organizations should stay on top of new advancements in foundation models and invest enough resources and expertise to leverage these technologies effectively for their specific industry.

5.4 H4: Industry-specific opportunities

The fourth hypothesis was: Opportunities arising from TL approaches are industry-specific. TL can enhance efficiency, accuracy, and innovation in various domains. Identifying and capitalizing on these opportunities require domain-specific expertise and tailored strategies.

Industry-specific opportunities for TL exist. However, industries mostly share a common set of general TL opportunities. TL has been proven to enhance efficiency, accuracy and innovation in all researched industries, not only for a specific industry. TL for instance reduces the need for computing power and data while simultaneously improving the development time and complexity of AI development. This is beneficial for for instance the healthcare industry, because it allows for the development of more accurate diagnostic tools and personalized treatment plans, ultimately improving patient outcomes and reducing healthcare costs. Additionally, TL can facilitate the integration of EHRs and medical imaging data, enabling more comprehensive and data-driven approaches to healthcare delivery. But it is also beneficial for the automotive and finance industry as it can enable the development of better smart manufacturing processes and streamline processes such as fraud detection, risk assessment, and customer relationship. TL can bring a lot of efficiency gains to all industries. All opportunities that arise from TL can mostly be derived from the earlier mentioned reasons for why organizations use TL. Therefore, this hypothesis can be partially accepted.

As show in the above explained examples, industry-specific opportunities are mostly use-case based. TL can for instance be used for easier extension of fraud-detection algorithms to different countries. This is an opportunity specific for the finance industry, but is use-cased based. The use-case is specific for one industry. However, if a use-case can be altered to fit another industry, it would not eliminate the use of TL. Another example is the expected impact of federated TL applications for the healthcare industry. This could create more impact AI applications while keeping the privacy of patients intact.

Organizations should acknowledge the general opportunities for the use of TL and take this with them in their AI strategies. TL offers opportunities that make AI development easier and more efficient. By embracing these advantages, organizations can significantly improve their overall AI strategies and drive meaningful advancements in their respective domains. It is recommended that industries look beyond their own borders and use-cases. For example, a TL based computer vision solution works in the automotive sector, it could also work in the healthcare industry. Being able to identify solutions in different industries could lead to cross-pollination of ideas and innovation, accelerating the pace of technological advancement and fostering collaboration across diverse sectors.

5.5 The future of the TL landscape

After a thorough examination of the hypotheses, an exploration of the future of TL in organizations will be conducted. What are trends regarding TL? How will advancements in NLP and computer vision shape the future of TL applications?

Organizations from all industries see big opportunities for further expansion of TL in NLP applications. It is expected that this will be the most important application for TL in the upcoming future. With the rise of foundation models and readily available LLM's such as GPT, it will probably become easier for organizations to apply and scale NLP-based TL solutions. This trend opens up possibilities for a wide range of applications, including automated text generation, sentiment analysis, language translation, and conversational AI. All applications that were first only accessible to organizations who had the ability to train LLM's themselves. By applying TL in collaboration with pre-trained LLM's. It is expected that the whole process of finetuning a model based on your own data will become easier in the future. OpenAI is already offering solutions for this, but platforms and low/no-code solutions for the fine-tuning of models are expected to arise.

Similarly, more advancements in computer vision TL technology too allow for more innovation in all industries. When the development of more accurate and efficient vision models becomes easier, this will enable organizations to deploy more sophisticated computer vision solutions for object detection, image classification, anomaly detection, and augmented reality. These are all applications suitable for the automotive, finance, and healthcare industry.

However, alongside these technological advancements, organizations must also navigate the ethical and regulatory considerations surrounding TL adoption. As AI technologies become more pervasive in society, there is increasing scrutiny on issues such as bias, fairness, transparency, and data privacy. Compliance with ethical AI principles and regulatory frameworks, such as GDPR and CCPA, will be paramount to ensure responsible and ethical use of TL in organizations. Organizations are recommended to develop suitable AI strategies which take all sides of TL into account.

5.6 Limitations and Recommendations

It is important to note that while federated TL emerged frequently in discussions, its focus on information management and regulatory considerations fell outside the scope of this research. More research on the relation between TL and federated learning would be recommended to examine if this really is a viable solution for for instance the healthcare industry.

Another recommendation for future studies would involve conducting a more extensive survey research. This exploratory research provided valuable insights, but a larger sample size is necessary for a comprehensive understanding of the state of TL implementation in different industries. Together with challenges in securing an adequate number of respondents stemmed because of the nature of TL, more in depth knowledge could be acquired by conducting an extensive survey.

Although this thesis covers various industries, each sector warrants further in-depth investigation. Future studies could delve deeper into the status of TL implementation within a single industry. More focus on one specific industry could identify more industry specific challenges, opportunities and reasons for the use TL.

Finally, there seems to be a big movement for TL towards the fine-tuning of foundation models. For a better understanding on this subject more research specialized on TL in relation to NLP and foundation models is recommended.

Chapter 6

Conclusion

The main objective of this research was to determine the current state of TL adoption in different industries and identify challenges and opportunities. Based on a qualitative approach, the following research question was addressed, supported by four hypotheses:

How is Transfer Learning currently being adopted across different industries and what are challenges and opportunities based on the identified TL approaches?

The research findings show the variability in TL adoption across industries, with some sectors demonstrating greater receptiveness than others. While TL proves to be applicable across all sectors, the extent of adoption varies, influenced by factors such as industry-specific needs, regulations, and mindset. Notably, the healthcare industry emerges as less receptive, primarily due to regulatory constraints and a mindset that is less favourable for innovation. In conclusion, TL is implemented reactively in organizations, based on specific use-cases. All capabilities of AI are suitable for the use of TL, but it seems to be most prevalent in use-cases for computer vision and NLP.

Multiple challenges make successful implementation of TL difficult for organizations, these include legal and regulatory complexities, possible biases of external models, technological hurdles and an overall mindset which facilitates the adoption of new technologies. These challenges are not unique to TL and can possibly be applied to other new technologies as well. However, they underscore the need for tailored strategies and expertise around TL to overcome them.

Despite the existence of industry-specific nuances, common opportunities that TL can bring organizations have been identified. These are mostly focused on enhanced efficiency and easier development of AI applications. Organizations can for example choose a TL solution because it can be developed with a lesser data need or overall monetary investment. TL's ability to streamline AI development processes and its potential to address various industry challenges underscore why it should have a position in every organizational AI strategy. Not only in a reactive, but in a proactive way as well.

This thesis has tried to grasp a view on TL adoption and its implications. However, it is crucial to address certain limitations. Future research should explore the relationship between TL and federated learning, conduct more extensive surveys focused on single industries, and delve deeper into specific industry challenges and opportunities. In addition, now that challenges and opportunities for organizations have been identified, the next we could focus on is how to overcome these challenges. Finally, specialized research on TL in relation to NLP and foundation models could provide valuable insights into emerging trends and applications.

In conclusion, this study offers valuable insights into the broad landscape of TL adoption, challenges, and opportunities across industries. By addressing the identified challenges and leveraging the opportunities presented by TL, organizations can drive innovation, efficiency, and competitiveness in their respective domains. Proven is that organization from all industries apply TL, but it still has a lot more to offer, which organizations should act upon.

Bibliography

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al. (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Amazon. (2023). Data labeling - Amazon sagemaker ground truth- amazon web services [[Online; accessed 2023-10-18]].

Anacleto, J. C., de Carvalho, A. F. P., Pereira, E. N., Ferreira, A. M., & Carlos, A. J. (2008). Machines with good sense: How can computers become capable of sensible reasoning? *Artificial Intelligence in Theory and Practice II: IFIP 20 th World Computer Congress, TC 12: IFIP AI 2008 Stream, September 7-10, 2008, Milano, Italy 2*, 195–204.

Antunes, M., Bergasa, L. M., Araluce, J., Gutiérrez, R., Arango, J. F., & Ocaña, M. (2022). Including transfer learning and synthetic data in a training process of a 2d object detector for autonomous driving. *Iberian Robotics conference*, 465–478.

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models.

Asawa, C., Elamri, C., & Pan, D. (2017). Using transfer learning between games to improve deep reinforcement learning performance and stability. https://api.semanticscholar.org/CorpusID:11306158

Bansal, M. A., Sharma, D. R., & Kathuria, D. M. (2022). A systematic review on data scarcity problem in deep learning: Solution and applications. *ACM Computing Surveys (CSUR)*, *54*(10s), 1–29.

Basias, N., & Pollalis, Y. (2018). Quantitative and qualitative research in business & technology: Justifying a suitable research methodology. *Review of Integrative Business and Economics Research*, *7*, 91–105.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., et al. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Bray, C. W. (1928). Transfer of learning. *Journal of Experimental Psychology*, *11*(6), 443.

Chen, Y., Wang, J., Yu, C., Gao, W., & Qin, X. (2021). Fedhealth: A federated transfer learning framework for wearable healthcare.

Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2017). Transfer learning for music classification and regression tasks. *arXiv preprint arXiv:1703.09179*.

Chui, M., Hall, B., Mayhew, H., Singla, A., & Sukharevsky, A. (2022). The state of ai in 2022—and a half decade in review. https://www.mckinsey.com/

capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review
Curreri, F., Patanè, L., & Xibilia, M. G. (2021). Soft sensor transferability: A survey. *Applied Sciences*, *11*(16), 7710.
Das, N., Bodapati, S., Sunkara, M., Srinivasan, S., & Chau, D. H. (2021). Best of both worlds: Robust accented speech recognition with adversarial transfer learning. *arXiv preprint arXiv:2103.05834*.
Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future healthcare journal*, *6*(2), 94.
den Hamer, P. (2023). Hype cycle for generative ai. *Gartner*.
ETO. (2023, September). Country activity tracker (cat): Artificial intelligence. https://www.bibme.org/bibliographies/c431772e-7faa-4643-9cc4-953f3dacc88d
Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *Bmj*, *323*(7321), 1103–1105.
Fang, X. (2021). Making recommendations using transfer learning. *Neural Computing and Applications*, *33*(15), 9663–9676.
Gartner. (2023). Gartner hype cycle research methodology [[Online; accessed 2023-11-22]].
Gautam, S., & Kumar, A. (2022). Automatic traffic light detection for self-driving cars using transfer learning, 597–606.
Google. (2023). Vertex AI pricing [[Online; accessed 2023-10-18]].
Hammarberg, K., Kirkman, M., & de Lacey, S. (2016). Qualitative research methods: When to use them and how to judge them. *Human reproduction*, *31*(3), 498–501.
Han, S.-H., Kim, K. W., Kim, S., & Youn, Y. C. (2018). Artificial neural network: Understanding the basic concepts without mathematics. *Dementia and Neurocognitive Disorders*, *17*(3), 83–89.
He, H., & Wu, D. (2019). Transfer learning for brain–computer interfaces: A euclidean space data alignment approach. *IEEE Transactions on Biomedical Engineering*, *67*(2), 399–410.
He, Q.-Q., Pang, P. C.-I., & Si, Y.-W. (2019). Transfer learning for financial time series forecasting. *PRICAI 2019: Trends in Artificial Intelligence: 16th Pacific Rim International Conference on Artificial Intelligence, Cuvu, Yanuca Island, Fiji, August 26–30, 2019, Proceedings, Part II 16*, 24–36.
Houlsby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., De Laroussilhe, Q., Gesmundo, A., Attariyan, M., & Gelly, S. (2019). Parameter-efficient transfer learning for nlp. *International Conference on Machine Learning*, 2790–2799.
Hua, J., Zeng, L., Li, G., & Ju, Z. (2021). Learning for a robot: Deep reinforcement learning, imitation learning, transfer learning. *Sensors*, *21*(4), 1278.
Iman, M., Arabnia, H. R., & Rasheed, K. (2023). A review of deep transfer learning and recent advancements. *Technologies*, *11*(2), 40.
Kheder, M. Q., & Mohammed, A. A. (2023). Transfer learning based traffic light detection and recognition using cnn inception-v3 model. *Iraqi Journal of Science*, 6258–6275.
Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia tools and applications*, *82*(3), 3713–3744.

Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E., & Ganslandt, T. (2022). Transfer learning for medical image classification: A literature review. *BMC medical imaging*, *22*(1), 69.

Kim, J., & Park, C. (2017). End-to-end ego lane estimation based on sequential transfer learning for self-driving cars. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 30–38.

Laurer, M., van Atteveldt, W., Casas, A., & Welbers, K. (2022). Less annotating, more classifying: Addressing the data scarcity issue of supervised machine learning with deep transfer learning and bert-nli. *Political Analysis*, 1–33.

Leiter, C., Zhang, R., Chen, Y., Belouadi, J., Larionov, D., Fresen, V., & Eger, S. (2023). Chatgpt: A meta-analysis after 2.5 months.

Leung, L. (2015). Validity, reliability, and generalizability in qualitative research. *Journal of family medicine and primary care*, *4*(3), 324.

Li, X., Chan, S., Zhu, X., Pei, Y., Ma, Z., Liu, X., & Shah, S. (2023). Are chatgpt and gpt-4 general-purpose solvers for financial text analytics? a study on several typical tasks.

Li, Z., Cheng, X., Peng, X. B., Abbeel, P., Levine, S., Berseth, G., & Sreenath, K. (2021). Reinforcement learning for robust parameterized locomotion control of bipedal robots. *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2811–2817.

Liang, Y., Monteiro, S. T., & Saber, E. S. (2016). Transfer learning for high resolution aerial image classification. *2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 1–8.

Lins, S., Pandl, K. D., Teigeler, H., Thiebes, S., Bayer, C., & Sunyaev, A. (2021). Artificial intelligence as a service: Classification and research directions. *Business & Information Systems Engineering*, *63*, 441–456.

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). Artificial intelligence index report 2023.

Myers, B. A. (1998). A brief history of human-computer interaction technology. *interactions*, *5*(2), 44–54.

Nguyen, D., Sridharan, S., Nguyen, D. T., Denman, S., Dean, D., & Fookes, C. (2020). Meta transfer learning for emotion recognition.

Niu, S., Liu, Y., Wang, J., & Song, H. (2020). A decade survey of transfer learning (2010–2020). *IEEE Transactions on Artificial Intelligence*, *1*(2), 151–166.

OpenAI. (2023). Openai platform [[Online; accessed 2023-10-25]].

Oseni, A., Moustafa, N., Janicke, H., Liu, P., Tari, Z., & Vasilakos, A. (2021). Security and privacy for artificial intelligence: Opportunities and challenges.

Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, *22*(10), 1345–1359.

Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M., & Dean, J. (2021). Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*.

Politou, E., Alepis, E., Virvou, M., & Patsakis, C. (2022). *Privacy and data protection challenges in the distributed era* (Vol. 26). Springer.

Rai, A., Constantinides, P., & Sarker, S. (2019). Next generation digital platforms: Toward human-ai hybrids. *Mis Quarterly*, *43*(1), iii–ix.

Ray, P. P. (2023). Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*.

Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative research in psychology*, *11*(1), 25–41.

Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., & Hadsell, R. (2016). Progressive neural networks. *arXiv preprint arXiv:1606.04671*.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & quantity*, *52*, 1893–1907.

Steinert, M., & Leifer, L. (2010). Scrutinizing gartner's hype cycle approach. *Picmet 2010 technology management for global economic growth*, 1–13.

Suryanto, H., Mahidadia, A., Bain, M., Guan, C., & Guan, A. (2022). Credit risk modeling using transfer learning and domain adaptation. *Frontiers in Artificial Intelligence*, *5*, 868232.

Tamburrini, G. (2022). The ai carbon footprint and responsibilities of ai scientists. *Philosophies*, *7*(1), 4.

Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., & Elger, G. (2021). Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability engineering & system safety*, *215*, 107864.

Thornberg, R., Charmaz, K., et al. (2014). Grounded theory and theoretical coding. *The SAGE handbook of qualitative data analysis*, *5*(2014), 153–69.

Topaloglu, M. Y., Morrell, E. M., Rajendran, S., & Topaloglu, U. (2021). In the pursuit of privacy: The promises and predicaments of federated learning in healthcare. *Frontiers in Artificial Intelligence*, *4*, 746497.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Verdegem, P. (2022). Dismantling ai capitalism: The commons as an alternative to the power concentration of big tech. *AI & society*, 1–11.

Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., et al. (2018). Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, *2018*.

Wang, M., & Yang, H. (2021). Research on personal credit risk assessment model based on instance-based transfer learning. *Intelligence Science III: 4th IFIP TC 12 International Conference, ICIS 2020, Durgapur, India, February 24–27, 2021, Revised Selected Papers 4*, 159–169.

Wang, Y., Gu, Q., & Brown, D. (2019). Differentially private hypothesis transfer learning. *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10–14, 2018, Proceedings, Part II 18*, 811–826.

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, *3*(1), 1–40.

Williams, C., et al. (2007). Research methods. *Journal of Business & Economics Research (JBER)*, *5*(3).
Woodward, K., Kanjo, E., Brown, D. J., & McGinnity, T. (2020). On-device transfer learning for personalising psychological stress modelling using a convolutional neural network. *arXiv preprint arXiv:2004.01603*.
Xie, M., Jean, N., Burke, M., Lobell, D., & Ermon, S. (2016). Transfer learning from deep features for remote sensing and poverty mapping.
Xu, R., Baracaldo, N., Zhou, Y., Anwar, A., Kadhe, S., & Ludwig, H. (2022). Detrust-fl: Privacy-preserving federated learning in decentralized trust setting.
Yoon, S., Yun, H., Kim, Y., Park, G.-t., & Jung, K. (2017). Efficient transfer learning schemes for personalized language modeling using recurrent neural network. *arXiv preprint arXiv:1701.03578*.
Zhang, L., & Gao, X. (2022). Transfer adaptation learning: A decade survey. *IEEE Transactions on Neural Networks and Learning Systems*.
Zhang, Y., Niu, S., Qiu, Z., Wei, Y., Zhao, P., Yao, J., Huang, J., Wu, Q., & Tan, M. (2020). Covid-da: Deep domain adaptation from typical pneumonia to covid-19.
Zhao, P., Hoi, S. C., Wang, J., & Li, B. (2014). Online transfer learning. *Artificial intelligence*, *216*, 76–102.
Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., Zhang, K., Ji, C., Yan, Q., He, L., et al. (2023). A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. *arXiv preprint arXiv:2302.09419*.
Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76.

Appendix A

Interview questions

Introduction

Personal introduction and introduction to research
Consent for recording and anonymity of interviewee and organization

- Could you introduce yourself?
  - Background
  - Experience
  - Job title
  - Company size and industry
- What is your background related to the implementation or oversight of AI?

General questions

- How do you define AI?
- How are you using machine learning or AI in work?
  - Personal use
  - Projects
- Do you know what the general concept of Transfer Learning in AI entails?
  - Could you explain this?

Definition Transfer Learning (we use this for the rest of the interview):
Transfer Learning can be described in AI as the concept of the improvement of learning in a new task through the transfer of knowledge from a related task

- Can you provide an example of how Transfer Learning is currently being used in your industry?
  - Why are you using it this way?
- What could be other ways of how TL can be used in your industry?

    - Why is TL the preferred method?

- How do you define the scope and objectives of Transfer Learning within your industry/organization?
    - Which aspects of the organization are utilizing TL?
    - What are primary TL goals?

- What are the primary data types used for Transfer Learning in your industry?
    - Do you also use open data, why (not)?
    - Compared to traditional machine learning models, how does the data acquisition process differ in scale and complexity for transfer learning?
    - Can you elaborate on the trade-off between the effort in labeling data and the efficiency gains in transfer learning models?
    - Which data sources are not suitable for TL?
    - How do you address data privacy and security concerns?

- Does your organization or industry have rules on which models can be used for transfer learning?

- I have identified 5 reasons for the use of TL in organizations, could you please rank them based on your experience on which are most decisive?
    - Is there a reason I missed?
    - Why do you rank them this way?
    - If we look at the future, are these reasons the same? (considering legislation, availability of more data and computing power)

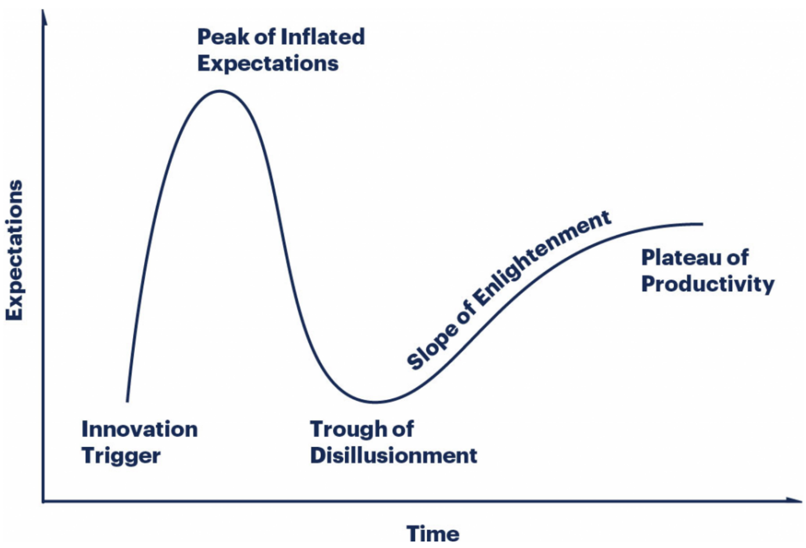
Reason	Now	Future
Data scarcity		
Lack of computational power		
Maintaining privacy		
Lesser environmental impact		
Personalized AI		

- Transfer Learning can be categorized in multiple ways: inductive, transductive and unsupervised. Which method are you using the most?
    - Could you explain why?

Definition Deep Transfer Learning

Deep transfer learning is a machine learning technique where a model developed for a specific task is reused as the starting point for a model on a second task, leveraging deep learning architectures. Examples of this are the use of pre-trained (foundation) models with freezing and retraining layers.

- An extension to Transfer Learning is Deep Transfer Learning, how are organizations in your industry using this?
  - Are you developing models yourself or using platforms? (e.g. OpenAI)
  - Do you see a trend towards Deep Transfer Learning or the use of foundation models in the industry compared to traditional transfer learning?
- What do you consider the biggest challenges in adopting and scaling Transfer Learning within your industry?
  - No knowledge of TL available, no data access, privacy concerns, homogeneity concerns
  - How do you plan to overcome these challenges?
- What do you perceive as the most significant opportunities for Transfer Learning in your industry in the next 5-10 years?
- Where would you put Transfer Learning on the technology hype cycle? Can you explain why



Appendix B

Reasons for the use of TL

Automotive related interviews

Reason	Now	Future
Data scarcity	1	-
Lack of computational power	2	-
Maintaining privacy	4	0
Lesser environmental impact	5	+
Personalized AI	3	+

Table B.1: Automotive interview 1

Reason	Now	Future
Data scarcity	2	0
Lack of computational power	3	0
Maintaining privacy	4	+
Lesser environmental impact	5	+
Personalized AI	1	0

Table B.2: Automotive interview 2

Healthcare related interviews

Reason	Now	Future
Data scarcity	1	-
Lack of computational power	4	- (1)
Maintaining privacy	2	+ (2)
Lesser environmental impact	3	+ (3)
Personalized AI	5	+

Table B.3: Healthcare interview 1

Reason	Now	Future
Data scarcity	1	-
Lack of computational power	5	-
Maintaining privacy	2	+
Lesser environmental impact	3	+
Personalized AI	4	+

Table B.4: Healthcare interview 2

Reason	Now	Future
Data scarcity	1	- (4)
Lack of computational power	4	+ (3)
Maintaining privacy	2	+ (5)
Lesser environmental impact	3	+ (2)
Personalized AI	5	+ (1)

Table B.5: Healthcare interview 3

Finance related interviews

Reason	Now	Future
Data scarcity	1	-
Lack of computational power	2	-
Maintaining privacy	5	+
Lesser environmental impact	4	+
Personalized AI	3	+

Table B.6: Finance interview 1

Reason	Now	Future
Data scarcity	2	-
Lack of computational power	1	-
Maintaining privacy	4	+
Lesser environmental impact	3	+
Personalized AI	5	-

Table B.7: Finance interview 2

General interviews

Reason	Now	Future
Data scarcity	1	-
Lack of computational power	2	-
Maintaining privacy	5	+
Lesser environmental impact	3	+
Personalized AI	4	+

Table B.8: General interview 1

Reason	Now	Future
Data scarcity	1	-
Lack of computational power	2	+
Maintaining privacy	5	+
Lesser environmental impact	4	+
Personalized AI	3	+

Table B.9: General interview 2

Reason	Now	Future
Data scarcity	1	0
Lack of computational power	2	0
Maintaining privacy	4	+
Lesser environmental impact	3	+
Personalized AI	5	+

Table B.10: General interview 3

Reason	Now	Future
Data scarcity	1	+
Lack of computational power	2	+
Maintaining privacy	3	+
Lesser environmental impact	5	+
Personalized AI	4	0

Table B.11: General interview 4