Building a Credit Risk Model using Transfer Learning and Domain Adaptation

Hendra Suryanto,[1] Charles Guan,[1] Ada Guan,[1] Pengqian Li,[1]
Paul Compton,[2] Michael Bain,[2] Ghassan Beydoun[3]
[1]Rich Data Corporation
[2]University of New South Wales
[3]University of Technology Sydney
{hendra.suryanto, charles.guan, ada.guan, patrick.li}@richdataco.com,
{p.compton, m.bain}@unsw.edu.au, ghassan.beydoun@uts.edu.au

Abstract

Transfer learning has been successfully applied to the credit risk domain to predict the probability of default for "new to credit" individuals and small businesses. However when the source and target domains differ, we propose a domain adaptation approach to adjust the source domain features. We find that adaptation improves model accuracy in addition to the improvement by transfer learning. We propose and test a combined strategy of feature selection and an adaptation algorithm to convert values of source domain features to mimic target domain features. We find that transfer learning improves model accuracy by increasing the contribution of less predictive features. Although the percentage improvements are small, such improvements in real world lending would be of great economic importance. Our contribution also includes a strategy to choose features for adaptation and an algorithm to adapt values of these features.

Introduction

Globally in 2014, 42% of all adults reported borrowing in the past 12 months (excluding credit cards). In developing economies three times as many adults borrowed from family or friends as from a financial institution. Borrowing from an institution has benefits over borrowing from family or friends, by providing access to sufficient funds and likely better credit terms under regulation (World Bank 2017). Access to formal credit has become an issue for young adults in developed countries too. Bankrate's survey found that 58% of millennials (born between 1981 and 1996) in the United States have been denied at least one financial product because of their credit score (BankRate 2019). As well, fintech-based financing products, such as Alipay, Affirm, Klarna, Paypal Credit and Afterpay, have become popular with millennials and Generation Z (born between 1997 and 2010). Can we leverage this "alternative lending" data to improve prediction of credit behaviour and hence access to credit for people with limited traditional credit history?

Transfer learning can be the bridge linking alternative lending data and traditional credit history assessment, e.g., credit bureau scores. Suryanto et al. (2019) (Suryanto et al. 2019) showed transfer learning improved the accuracy of credit scoring. To adopt this approach in the real world, two questions need to be answered.

The first is to explain transferred models. Many jurisdictions require credit decisions to be explained for anti-discrimination and human rights purposes. For example, the General Data Protection Regulation (GDPR) in the European Union requires "meaningful information about the logic involved" in automated decisions, providing an explanation that enables a data subject to exercise their rights under GDPR and human rights law (Selbst and Powles 2017). SHAP (Lundberg and Lee 2017) is one of the most popular methods for explaining machine learned models. In this paper we apply SHAP to analyse the contribution of features and the impact of transfer.

The second question is how to handle the difference in features between source and target domains. For instance, a source domain could be for a small short-term alternative loan, but the target domain may be for large and long-term instalment loans. Key features, such as loan amount, loan terms, interest cover, etc. can differ between these domains. We could use the progressive shifting contribution network proposed in (Suryanto et al. 2019) that combines source and target domain feature learning to improve model accuracy, but a key question that remains is: can we adapt the features *before* transfer learning to get more accurate models?

In this paper we develop an approach to this question based on three approaches. First, we use a Kolmogorov–Smirnov (KS) test to quantify the difference between source and target domains, and use domain adaptation to treat only features that differ substantially between the domains before training. Second, after we find candidate features to be adapted, based on their KS differences, we include other features that are highly correlated with the candidate features and test the accuracy of models adapting these feature combinations. Finally, we exclude from adaptation features related to a borrower's credit history where the adaptation would incorrectly impact the classification. The rest of the paper is organised as follows. In Section we describe key aspects of the problem and in Section related work. In Section we elaborate details of the data and methods; in Section , the experimental results and in Sections and discussion and conclusions.

Credit Scoring and Decisioning

A lender's goal is to maximise the risk adjusted return within their risk appetite. Accurately assessing credit risk is key to balancing risk and return. The concept of Expected Loss (EL) is commonly used to measure credit risk. EL is mainly determined by the Probability of Default (PD), Exposure at Default and Loss Given Default. The key prediction model is a credit scoring model, which calculates the PD of a loan or loan application. Inputs to a credit scoring model are attributes of the person or entity applying for the loan, such as credit history, credit bureau score and employment, and requested loan attributes, such as loan amount and term.

Lenders then use the credit score and other decision rules to decide whether to approve or decline a loan application, and for those approved what to offer in credit terms. Decision rules typically include: eligibility, e.g., age limit, residing jurisdiction; expert assessment of risks, e.g., manual reviews and override; credit scoring and rating, i.e. mapping the credit score to different credit ratings; and a decision table or scale, e.g., decline under a certain rating level, or set the maximum loan amount at certain ratings.

In this paper we use a score from 0 to 1 for PD models, which is an estimated probability of default, calibrated using test data. Our focus is on using transfer learning to predict PD, so we measure the accuracy of our credit scoring models using the Area Under Receiver Operating Curve (AUC). This metric is used to directly assess model accuracy, based on PD, without needing to convert the PD into a binary "yes" or "no". The quality of binary classifications depends not only on the PD model, but also on decision rules such as those above.

Related Work

Transfer learning and domain adaptation are mostly applied in computer vision (Wang and Deng 2018), speech recognition (Deng et al. 2013), and natural language processing (Mou et al. 2016). Transfer learning has also been proposed to improve reinforcement learning in the Atari game domain. Rusu et al. (2016) (Rusu et al. 2016) presented the Progressive Network, a transfer learning approach based on a neural network where the network was initially trained using source domain data. Next, one or more of the last layers of the network were retrained using target domain data (Rusu et al. 2016). Using a similar approach, Suryanto et al. (Suryanto et al. 2019) proposed transfer learning based on the progressive network configuration, applied to credit risk where the contribution of the source and the target domains can be shifted to optimize the model performance. There have been other recent studies applying transfer learning in the credit risk domain, mostly for credit scoring rather than credit decisioning (Beninel, Waad, and Mufti 2012), (Stamate, Magoulas, and Thomas 2015), (Suryanto and Compton 2004).

While the terms "transfer learning" and "domain adaptation" have been used interchangeably, we use transfer learning when the focus is the modelling configuration, and domain adaptation when the focus is on transforming the data. There are only a few published studies on domain adaptation for credit risk, e.g., Huang et al. (2018) proposed domain adaptation for transforming the data distribution (Huang and Chen 2018). In other domains approaches such as Balanced Distribution Adaptation (Wang et al. 2017) and adapting without target label have been used (Kouw and Loog 2019; Zhang, Li, and Ogunbona 2018; Huang and Chen 2018).

In this paper we adopt a Progressive Network configuration for transfer learning, similar to Rusu et al. (Rusu et al. 2016). The contribution of our paper is a strategy to apply domain adaptation to the source data when target data with labels is limited, and to apply both domain adaptation and transfer learning to credit risk.

Data and Methods

Data

In this paper, we used data from the "lendingclub.com" dataset[1] to illustrate our approach. We used the purpose of the loans to define different domains. Loans for different purposes have different default rates and different loan parameters, such as the typical loan amount and the terms. The experiments in this paper were based on data for three different purposes:

- Data where the purpose of the loan was credit card and debt consolidation, which is referred to as CD.
- Data where the purpose of the loan was medical, referred as MD.
- Data where the purpose of the loan was small business lending, referred as SB.

In this empirical study, we used Lending Club (LC) data between 2007 and 2011, the early period of the Lending Club, to mimic a lender starting to offer new loan products to new customer segments. We used the CD dataset as the source domain, as it had sufficient instances, and the MD and SB datasets as target domains for transfer learning. Domain details are illustrated in Table 1.

Table 1: Loan domains by purpose

No	Domain	Number of Rows	Default Rate
1	CD	28,813	14.03%
2	MD	695	15.25%
3	SB	1,813	26.16%

To predict loan outcomes, we selected the 12 input features listed in Table 2. The PD model predicts whether loans should be classified as default or not. We use loan status to determine this outcome, as shown in Table 3. Based on our experience in credit risk, we only include loans with the loan status of Charged off or Late (31-120 days) as default, i.e., bad loans, and Fully Paid as good loans. We exclude current loans (not due yet), and loans less than 30 days late, which will normally be repaid but for which there are no results yet.

[1] See https://www.lendingclub.com/info/download-data.action

Table 2: Input features

No	Short Name	Feature Name & Description
1	term_36m	Term 36 month; The 36-month payment on the loan
2	term_60m	Term 60 month; The 60-month payment on the loan
3	grade_n	Grade; Lending Club (LC) assigned loan grade
4	sub_grade_n	Sub grade; LendingClub assigned loan subgrade
5	int_rate_n	Interest rate; Interest rate on the loan
6	revol_util_n	Revolving util. rate; Revolving line utilisation rate: the amount of credit relative to all available revolving credit
7	emp_length_n	Employment Length; Employment length in years: Values between 0 and 10 where 0 means less than one year and 10 means ten or more years.
8	dti_n	Debt to income ratio; The ratio of the borrower's total monthly debt payments on the total debt obligations, excluding mortgages and the requested LC loan, to the borrower's self-reported monthly income
9	installment_n	Installment; The monthly payment owed by the borrower if the loan is made
10	annual_inc_n	Annual income; The combined self-reported annual income provided by the co-borrowers during registration
11	loan_amnt_n	Loan amount; The listed amount of the loan applied for. If at some time, the credit department reduces the loan amount, this will be reflected in this value.
12	cover	Cover; A ratio calculated using the annual income on the loan amount (annual_inc_n/loan_amnt_n)

Transfer Learning

In this paper we use one of the neural network configurations proposed in (Suryanto et al. 2019) for transfer learning. The neural network comprises an input layer with 12 input nodes, aligned with 12 input features. The output layer consists of one output node. The output is a score between 0 and 1. This score is calibrated to be the probability of default (PD) as illustrated in Fig. 1. As our aim is to accurately predict defaults in the target domain, we first trained the model using source domain data, and then retrained the last layer with target domain data. We tested the accuracy of the transferred model using target domain data.

For comparison, we trained a "target model" using a similar neural network configuration with purely target domain data, and tested this on other target domain data. Suryanto et al. (2019) tested this configuration with state-of-the-art machine learning techniques for credit risk, e.g., gradient boosting machines, and the performance is equivalent (Suryanto et al. 2019). We used 10-fold cross validation, repeated 10 times using different random seeds, for all training and testing.

To answer the two questions on explainability and domain differences, we designed a set of experiments as described in the following sections.

Table 3: Outcome to predict: default or not

No	Loan Status	Description	Outcome
1	Charged off	The loan has not been paid	1
2	Fully Paid	The loan has been fully paid	0
3	Current	Payment is not due yet	excluded
4	In Grace Period	Payment is less than 16 days late	excluded
5	Late (16-30 days)	Payment is late between 16 and 30 days	excluded
6	Late (31-120 days)	Payment is late between 31 and 120 days	1

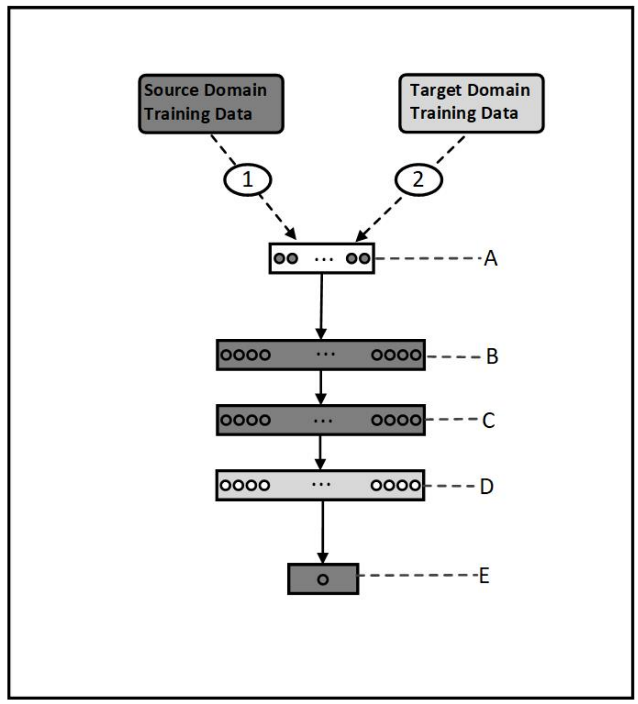

Figure 1: Transfer learning setup: layers B, C, D, and E are first trained using the Source Domain, then the last layer is retrained using the Target Domain; more precisely, the weights of the edges between layers D and E are retrained.

Domain Differences

To understand the differences between source and target domains, we use Kolmogorov–Smirnov (KS) tests to quantify the difference for each feature. The KS test can be used to compare two samples without making an assumption about the distribution of data. The null hypothesis is that the two samples, source and target data, come from the same distribution. The KS test produces a KS-statistic and p-value. The KS-statistic represents the maximum distance between the source data and the target data distributions. The p-value represents the significance level, e.g., less than 0.05. We used the maximum distance between the source data and the target data distribution curves (KS-statistic) to provide insights about the differences in features between these two domains.

Domain Adaptation

Domain adaptation aims to transform the source data distribution to be similar to the target data distribution. The intention is to use latent features constructed using source data to complement the target data. We propose the following approach to adapt the feature distribution of the source data to mimic the feature distribution of target data. For each feature, the adaptation steps are:

1. Group the source data and the target data in N quantiles, where N should be selected to ensure that we have sufficient data for each quantile, e.g., more than 50 samples. In this study, we selected $N = 10$, after experimenting with various N values.
2. For each corresponding source and target quantiles, calculate $scale$, then adapt/adjust the source feature values:

$$scale = \frac{(max(target\_value) - min(target\_value))}{(max(source\_value) - min(source\_value))} \tag{1}$$

$$offset = (source\_value - min(source\_value) * scale \tag{2}$$

$$adapted\_source\_value = min(target\_value) + offset \tag{3}$$

The adapted source features are used to initially train the neural network before the last layers are retrained using the target features.

Based on observation of explainer models and feature differences, we adapted different sets of features before training, and then trained and tested using the method described in Section on Transfer Learning. We then compared the performance of models (using AUC) with different adaptation sets, and transferred models without adaptation. We found that adapting all features significantly reduces accuracy, so we tried different combinations of features to adapt to find the most accurate adapted models.

Experiments and Results

Transfer Learning

Table 4 shows the AUC comparison for target and the transferred model. The accuracy of the transferred models was better than for the target models; AUC improved 0.042 or 7% for the MD domain, and 0.0224 or 3.6% for the SB domain, respectively. This is in line with the results of Suryanto el al. (Suryanto et al. 2019)

Table 4: Target model vs. transferred model

Domain & Experiment	AUC	Improvement	p-value
CD to MD; Training using Target only	0.5971 $\pm$0.08		
CD to MD; Training using Source then retraining the last layer using Target	0.6391 $\pm$0.09	0.0420 (7.0%)	<0.01
CD to SB; Training using Target only	0.6194 $\pm$0.05		
CD to SB; Training using Source then retraining the last layer using Target	0.6419 $\pm$0.05	0.0224 (3.6%)	<0.01

To understand the contribution of "cover", we calculated KS-statistics which represents the difference in value distribution for "cover" between source and target domains as shown in Figure 2 where source was CD and target was MD (KS-statistics: 0.2736) and Figure 3 where source was CD and target was SB (KS-statistics: 0.0585). The X-axis represents the value of "cover" and the Y-axis represents the number of loans. Further results are presented in following sections.

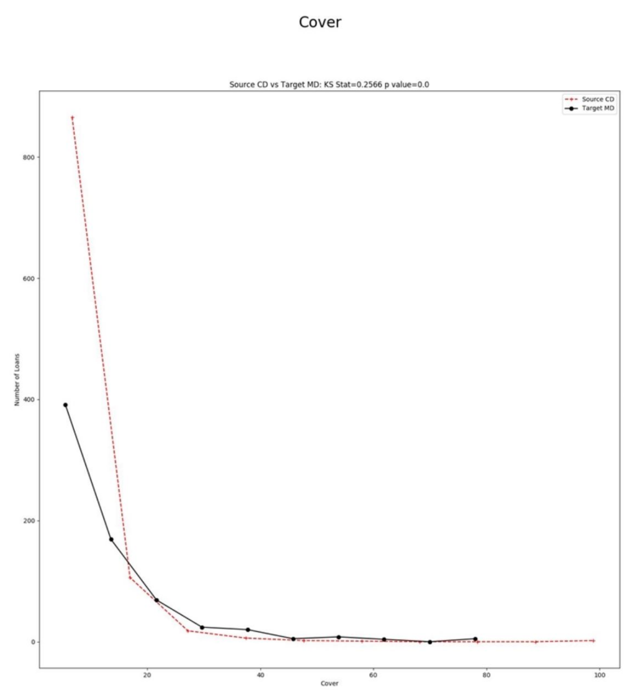

Figure 2: Distribution of cover: CD vs MD

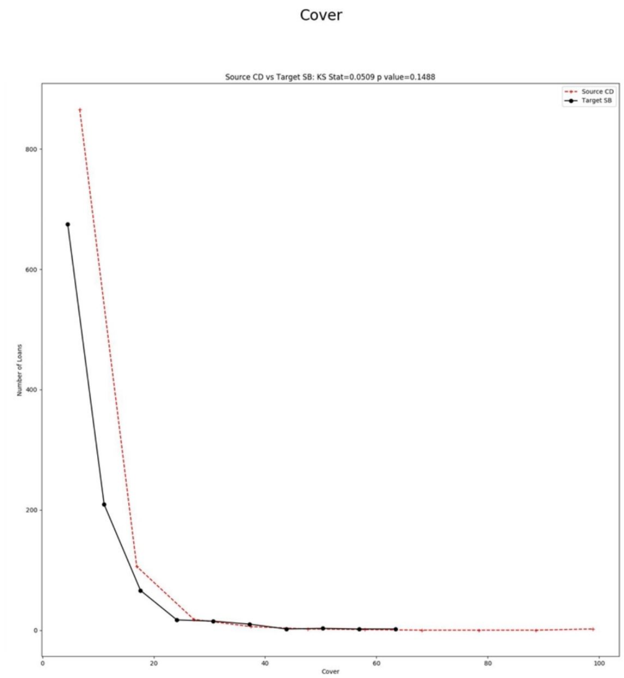

Figure 3: Distribution of cover: CD vs SB

Domain Difference

Table 5 lists KS-statistics for CD vs. MD as well as CD vs. SB. It shows that some features were very different between source and target domains, but some were similar. It also shows that different pairings of source and target domains had different patterns in feature differences. For example, "cover" was very different between CD and MD with a KS-statistic of 0.2729, but similar between CD and SB with a KS-statistic of 0.0502.

Table 5: KS of input features

		CD vs. MD		CD vs. SB	
No	Short Name	KS stats	KS p-value	KS stats	KS p-value
1	term_36m	0.0407	<0.24	0.0357	<0.20
2	term_60m	0.0407	<0.24	0.0357	<0.20
3	grade_n	0.0792	<0.01	0.0984	<0.01
4	sub_grade_n	0.0884	<0.01	0.1069	<0.01
5	int_rate_n	0.0941	<0.01	0.1033	<0.01
6	revol_util_n	**0.2248**	<0.01	**0.2292**	<0.01
7	emp_length_n	0.0242	<0.85	0.0749	<0.01
8	dti_n	**0.1502**	<0.01	**0.2295**	<0.01
9	installment_n	**0.3005**	<0.01	0.0671	<0.01
10	annual_inc_n	0.0670	<0.01	0.0906	<0.01
11	loan_amnt_n	**0.2899**	<0.01	0.0813	<0.01
12	cover	**0.2736**	<0.01	0.0585	<0.01

Domain Adaptation

To further understand the contribution of "cover", we tested our proposed domain adaptation function on "cover". Before we trained the transfer model on the source data, we adapted cover on source data to make it similar to the target data, and then applied the transfer learning technique to produce an "adapted" and transferred model. The AUC tests for these adapted and transferred models are listed in Table 6 where they are compared to the transferred model without adaptation. We have run paired t-tests to test the improvements shown in table 6, 7, 8; the improvements were all statistically significant with p-values <0.01. T-tests were appropriate because this data was normally distributed. Adapting cover works for CD to MD transfer with an AUC 0.01 (1.6%) higher than the transfer-only model, but AUC decreases for a CD to SB transfer.

Table 6: Adapted model vs. transferred model

Domain & Experiment	AUC	Improvement	p-value
CD to MD; Transfer only	0.6391 $\pm$0.0856		
CD to MD; Transfer with cover adapted	0.6491 $\pm$0.0824	0.0100 (1.6%)	<0.01
CD to SB; Transfer only	0.6419 $\pm$0.0509		
CD to SB; Transfer with cover adapted	0.6361 $\pm$0.0502	$-$0.0058 ($-$0.9%)	<0.01

We tested various permutations of features to adapt to find the most accurate model for the CD to MD transfer, and to establish an optimal strategy for seeking the most accurate adapted model. The experiments on the CD to MD transfer are listed in Table 7. Adapting all features, or adapting credit grade and related features, significantly reduced model accuracy, with AUC 0.1771 (27.7%) or 0.1756 (27.5%) lower than the transfer-only model, respectively. Adapting only features with a high KS-statistic (over 0.15), i.e., revolving utility, debt to income ratio, installment, loan amount, and cover, improved accuracy with AUC 0.0172 (2.7%) higher than the transfer-only model. Adding related features, i.e., annual income (annual_inc_n) – which is used to derive cover (a high KS feature), further improved accuracy, with AUC 0.0209 (3.3%) higher than the transfer-only model. Removing credit history features that are intrinsic to the borrower, i.e., revolving utility and debt to income ratio, produced an even more accurate model, with AUC 0.0257 (4.0%) higher than the transfer-only model.

Grade, sub-grade, revolving utility (revol_util_n), and debt to income ratio (dti_n) are features derived from credit history, which are intrinsic to the borrower and are usually highly correlated with the lending outcome, i.e., default or not. The interest rate in the lending club data is derived directly from grade and sub-grade, so we consider it as a credit history feature in our experiment.

Table 7: Adapted model vs. transferred model in CD to MD transfer

Experiment	AUC	Improvement	p-value
Transfer only	0.6391 ±0.0856		
Adapt all features	0.4620 ±0.3048	−0.1771 (−27.7%)	<0.01
Adapt credit grade and related features, i.e. grade, sub-grade, interest rate	0.4635 ±0.3052	−0.1756 (−27.5%)	<0.01
Adapt features with high KS, i.e. revolving utility, debt to income ratio, installment, loan amount and cover	0.6563 ±0.0806	0.0172 (2.7%)	<0.01
Adapt features with high KS and related features, i.e. revolving utility, debt to income ratio, installment, loan amount, cover and annual income	0.6600 ±0.07417	0.0209 (3.3%)	<0.01
Adapt features with high KS and related features less credit history features, i.e. installment, loan amount, cover and annual income	0.6649 ±0.0731	0.0257 (4.0%)	<0.01

The AUC comparison with the transfer-only model is shown in Table 8. Adapting all features, or credit grade related features, significantly reduced model accuracy, with AUC 0.123 (19.3%) or 0.1106 (17.2%) lower than the transfer-only model, respectively. We tested adaptation of the features that we adapted for the most accurate model of the CD to MD transfer, which have a low KS-statistic from CD and SB comparisons. This adapted model was slightly less accurate, with an AUC 0.0015 (0.2%) lower than the transfer-only model. Adapting features with a high KS-statistic, i.e., revolving utility and debt to income ratio, improved model accuracy slightly, with AUC 0.0018 (0.3%) higher than the transfer-only model. These two features do not have related features, and both were credit history features, so we could not improve accuracy further as we did with the CD to MD transfer.

Additionally, we investigated the explainability of the most accurate models using SHAP. Figures 4 and 5 show the feature contributions of the most accurate adapted models comparing to the source and target models. Through domain adaptation, the contribution of weak features increased in the most accurate adapted models. For the CD to MD transfer, the contribution of annual income, cover, installment and loan amount increased. For the CD to SB transfer, the contribution of annual income, term 36 months or 60 months, cover, employment length, installment and loan amount increased.

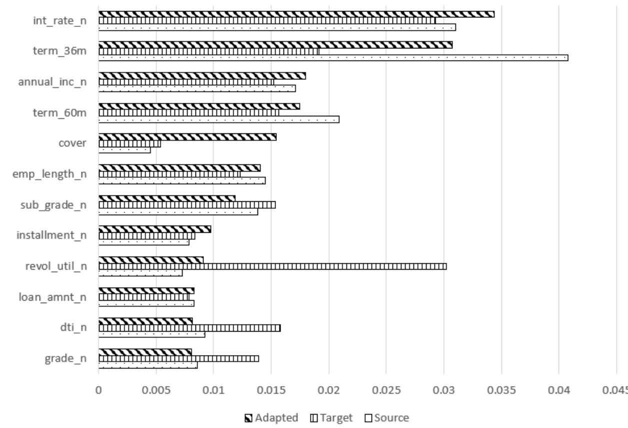

Figure 4: Feature contribution of the most accurate adapted model in CD to MD transfer

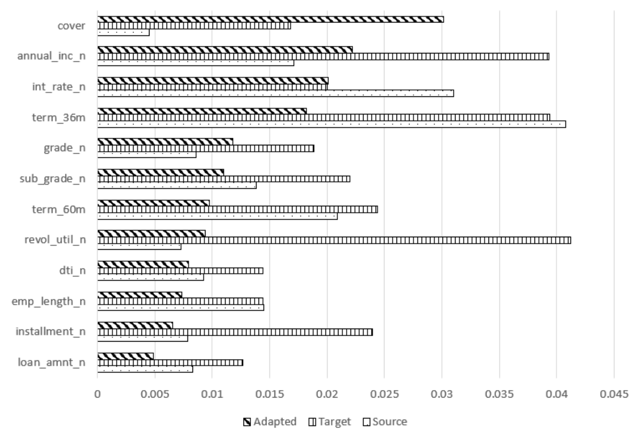

Figure 5: Feature contribution of the most accurate adapted model in CD to SB transfer

To evaluate the effectiveness of our adaptation approach we compared KS values before and after adaptation for the most accurate models, as shown in Table 9. The reduction in KS-statistics was between 44.8% and 90.3%, and for features with high KS-statistics (over 0.15) the reductions were all above 67.4%. Our adaptation approach successfully reduced the differences between the distribution of the source data and the target data.

Discussion

Transfer Learning improves model accuracy through generating intermediate features from the source domain to be selected for retraining on the target domain. This intermediate features generation concept is similar to "self taught learning" proposed by Raina et al. (2007) (Raina et al.

Table 8: Adapted model vs. transferred model in CD to SB transfer

Experiment	AUC	Improvement	p-value
Transfer only	0.6419 ±0.0509		
Adapt all features	0.5189 ±0.1666	-0.123 (-19.2%)	<0.01
Adapt credit grade and related features, i.e. grade, sub-grade, interest rate	0.5313 ±0.1624	-0.1106 (-17.2%)	<0.01
Adapt features used in CD to MD transfer, i.e. installment, loan amount, cover, annual income	0.6404 ±0.0475	-0.0015 (-0.2%)	<0.01
Adapt features with high KS, i.e. revolving utility and debt to income ratio	0.6437 ±0.0495	0.0018 (0.3%)	<0.01

2007), which constructed higher-level features using unlabelled data, except that in this paper we used labelled data from the source domain.

The contribution of a weak feature from the target domain increased because it was complemented by new intermediate features from the source domain. We tested an adaptation approach taking the outcome label into consideration. But this did not improve model accuracy. The reason was that the population of positive (outcome=1) cases was too small in the already small target dataset.

Adapting strong credit history features, such as grade and sub-grade, significantly reduced model accuracy, while removing features related to credit history from the adaptation list improved model accuracy. Adapting credit history related features *without consideration of the outcome label* generates unrealistic instances, e.g., changing a borrower's credit grade from high to low without adjusting the outcome from not default to default. These unrealistic instances can negatively impact model accuracy.

Conclusion

Domain adaptation with the right set of features further improved the accuracy of transfer learning models. However, adapting all features normally reduces model accuracy significantly. Reasons to select features to adapt include: differences in feature distribution between source and target domain, quantified by KS-statistics; relationships to already selected features; and domain specific knowledge, e.g., the credit history features intrinsic to the borrowers.

Through domain adaptation, the contribution of weaker features increased in the most accurate adapted models. An adaptation approach that significantly reduces KS-statistics has been critical in producing a successful domain adaptation algorithm.

Table 9: Kolmogorov-Smirnov test to compare source data and target data, before and after the source data is adapted, ACD is the abbreviation for Adapted Credit card and Debt consolidation data.

	CD to MD No adaptation		ACD to MD with adaptation		
Feature	KS-stats	p-value	KS-stats	p-value	Reduc-tion
installment	0.3005	<0.01	0.0293	<0.64	90.3%
annual_inc	0.0670	<0.01	0.0369	<0.34	44.8%
loan_amnt	0.2899	<0.01	0.0681	<0.01	76.5%
cover	0.2736	<0.01	0.0892	<0.01	67.4%
	CD to SB No adaptation		ACD to SB with adaptation		
Feature	KS-stats	p-value	KS-stats	p-value	Reduc-tion
revol_util	0.2292	<0.01	0.0536	<0.01	76.6%
dti	0.2295	<0.01	0.0301	<0.39	86.9%

For future work, the proposed strategy to select features for domain adaptation produces more accurate credit scoring models, but execution of the strategy requires human intervention in observing and applying domain knowledge. We will further explore methods to automate this selection strategy, so it can be a pre-processing step for fully automated transfer learning.

The use of alternatives to KS statistics to estimate the distance between distributions, such as KL-divergence, should be investigated. SHAP is an indirect method to understand the impact of latent intermediate features. Further study exploring and explaining latent intermediate features could improve our understanding of transfer learning and domain adaptation, and better meet transparency and compliance requirements.

Finally we note that although the significant improvements in accuracy demonstrated are small in terms of percentage improvements, such improvements in real world lending could be of substantial economic importance in reducing lenders' losses due to loan defaults.

References

BankRate. 2019. 58% of millennials have been denied at least one financial product because of their credit score.

Beninel, F.; Waad, B.; and Mufti, G. B. 2012. Transfer learning using logistic regression in credit scoring. *ArXiv* abs/1212.6167.

Deng, J.; Zhang, Z.; Marchi, E.; and Schuller, B. 2013. Sparse autoencoder-based feature transfer learning for speech emotion recognition. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 511–516.

Huang, J., and Chen, M. 2018. Domain adaptation approach for credit risk analysis. In *Proceedings of the 2018 International Conference on Software Engineering and Information*

Management, ICSIM2018, 104–107. New York, NY, USA: Association for Computing Machinery.

Kouw, W. M., and Loog, M. 2019. A review of domain adaptation without target labels. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1–1.

Lundberg, S., and Lee, S. 2017. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, 4765–4774.

Mou, L.; Meng, Z.; Yan, R.; Li, G.; Xu, Y.; Zhang, L.; and Jin, Z. 2016. How transferable are neural networks in nlp applications? In *EMNLP*.

Raina, R.; Battle, A.; Lee, H.; Packer, B.; and Ng, A. Y. 2007. Self-taught learning: Transfer learning from unlabeled data. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, 759–766. New York, NY, USA: Association for Computing Machinery.

Rusu, A. A.; Rabinowitz, N. C.; Desjardins, G.; Soyer, H.; Kirkpatrick, J.; Kavukcuoglu, K.; Pascanu, R.; and Hadsell, R. 2016. Progressive neural networks. *ArXiv* abs/1606.04671.

Selbst, A., and Powles, J. 2017. Meaningful information and the right to explanation. *International Data Privacy Law* 7(4):233–242.

Stamate, C.; Magoulas, G. D.; and Thomas, M. S. C. 2015. Transfer learning approach for financial applications. *ArXiv* abs/1509.02807.

Suryanto, H., and Compton, P. 2004. Invented predicates to reduce knowledge acquisition. In Motta, E.; Shadbolt, N. R.; Stutt, A.; and Gibbins, N., eds., *Engineering Knowledge in the Age of the Semantic Web*, 293–306. Berlin, Heidelberg: Springer Berlin Heidelberg.

Suryanto, H.; Guan, C.; Voumard, A.; and Beydoun, G. 2019. Transfer learning in credit risk. In *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. Springer.

Wang, M., and Deng, W. 2018. Deep visual domain adaptation: A survey. *Neurocomputing* 312:135–153.

Wang, J.; Chen, Y.; Hao, S.; Feng, W.; and Shen, Z. 2017. Balanced distribution adaptation for transfer learning. In *2017 IEEE International Conference on Data Mining (ICDM)*, 1129–1134. IEEE.

World Bank. 2017. Financial inclusion and inclusive growth.

Zhang, J.; Li, W.; and Ogunbona, P. 2018. Unsupervised domain adaptation: A multi-task learning-based method. *Knowledge-Based Systems*.