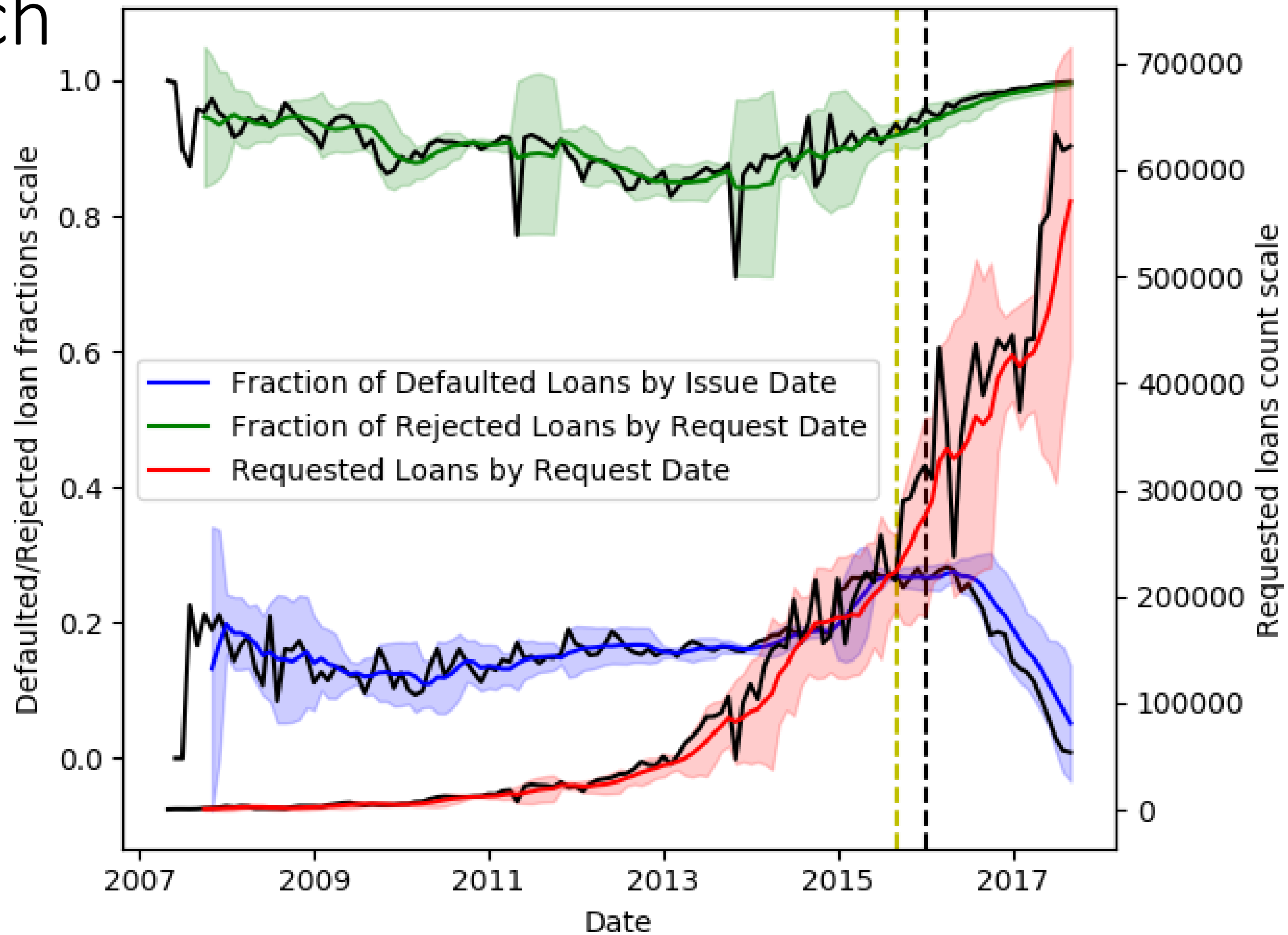
P2P lending: a case study

Loan screening and default prediction with Machine Learning and Deep Neural Networks

Tomaso Aste & J D Turiel

Aim of Research

- Automation of loan screening and acceptance through Machine Learning
- Accurate prediction of default risk through Big Data analytics and Machine Learning techniques
- P2P lending data investigated to understand limitations of default predictability



LendingClub

Data

- Open Dataset from The Lending Club (USA P2P lender)
- 15mln+ evaluated loans from 2007-2017
- Few available features for rejected loans
- 1mln issued loans from 2007-2017
- 150 available features for issued loans

Method: Two-Phase Model

First Phase
(Accepted/Rejected)

- Logistic Regression (Linear Kernel)
- Support Vector Machines (Gaussian Kernel)

Features:

- Debt to Income ratio
- Employment length
- Loan amount
- Loan Purpose

Second Phase
(Defaulted/Paid)

- Logistic Regression (Linear Kernel)
- Support Vector Machines (Gaussian Kernel)
- Linear Neural Network
- Deep Neural Network (***tanh*** node function)

Features:

- Loan amount
- Term
- Instalment
- Employment length
- Home ownership (owned, owned with a mortgage or rented)
- Verification status of the income
- Loan purpose
- Debt to Income ratio
- Earliest credit line in the record
- Number of open credit lines
- Number of derogatory public records
- Revolving line utilisation rate
- Total number of credit lines
- Number of mortgage credit lines
- Number of bankruptcies
- Logarithm of the applicant's annual income
- FICO score
- Logarithm of total credit revolving balance

Representative diagram outlining the two phases of the model with machine learning methods applied and features considered for each phase

Phase 1: Loan Selection

Loan Selection Results					
Model	Recall Train	AUC Test	Recall Macro Test	Recall Accepted Test	Recall Rejected Test
LR	79.8%	86.5%	77.4%	69.1%	85.7%
SVM	77.5%	-	75.2%	66.5%	84.0%

Results for the ML algorithms applied to the 1st model phase.

Parameters:
Imputation: Mean
Loss: log-likelihood
Optimiser: SGD
K-Fold Scoring: ROC-AUC
L2 Alpha: 10^{-3}

- Simple Logistic Regression model replicates analyst rejections with recall above 85%
- Target feature class imbalance in training set affects class scores. Would benefit from more training data
- Replicability of screening leads to more complex models applied to default prediction

Phase 2: Loan Default Prediction

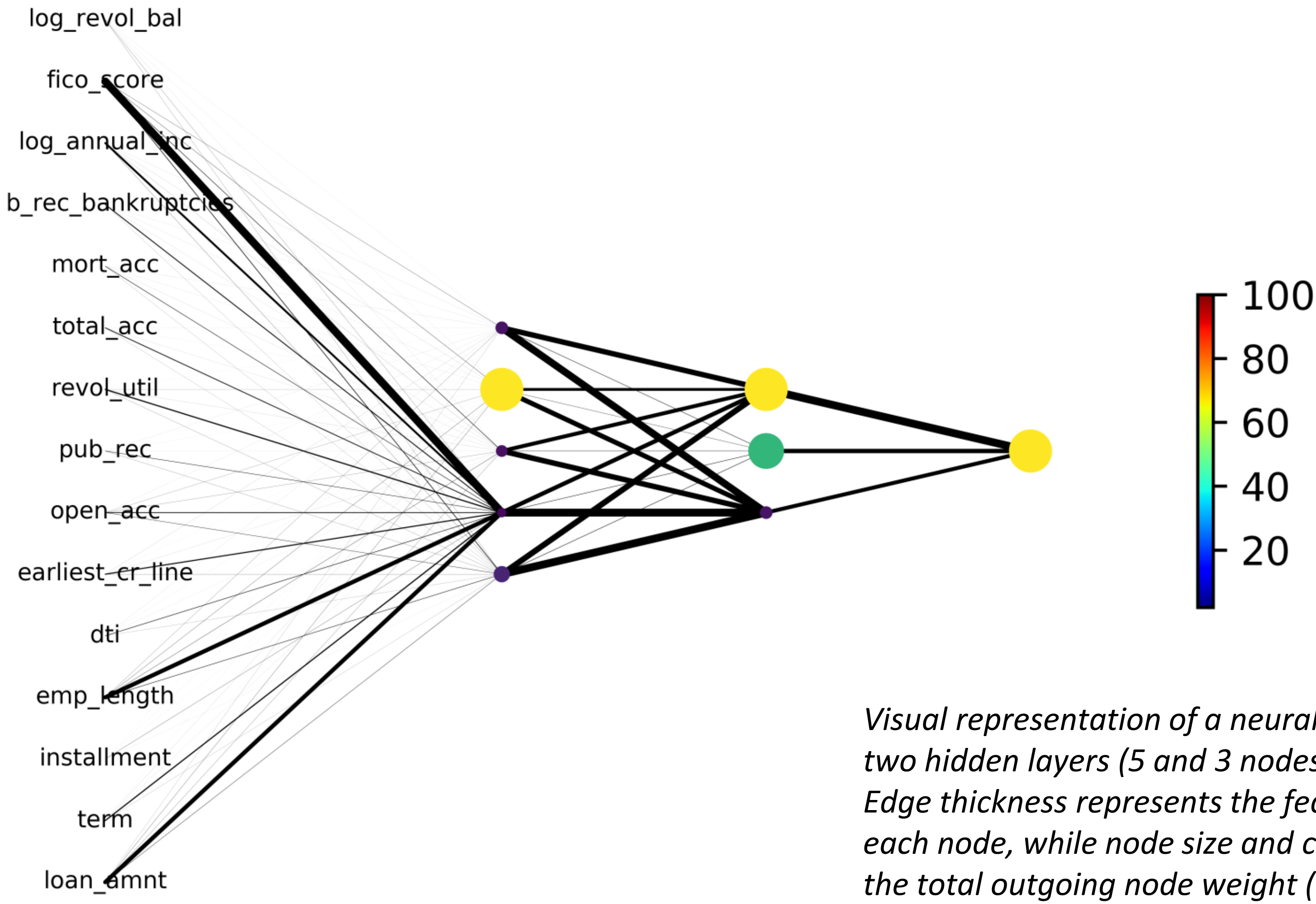
Loan Default Prediction Results					
Model	Recall Train	AUC Test	Recall Macro Test	Recall Default Test	Recall Paid Test
LR	64.3%	69.0%	63.7%	63.8%	63.6%
SVM	-	64.3%	62.15%	58.7%	65.6%
LNN [a]	-	67.8%	-	60.0%	-
LNN [b]	-	68.7%	-	62.7%	-
LNN [c]	-	69%	-	65%	-
DNN [d]	-	68%	-	67%	-
DNN [e]	71%	66%	-	75%	-
DNN [f]	68%	69%	-	72%	-

[a] LNN with numerical features only
[b] LNN with numerical and categorical features
[c] LNN with numerical and categorical features, L2 regularised
[d] DNN with arbitrary node numbers [20, 5]
[e] DNN with node numbers fine-tuned to [30, 1]
[f] DNN with node numbers fine-tuned to [5, 3]

Parameters:
Dropout: 20%
Imputation: Mean
Loss: log-likelihood
Optimiser: Adam
Activation: Sigmoid

Increasing complexity of the model captures the complex phenomenon of default, with DNNs outperforming in recall

Interpretability: Neural Network Visualisation



Visual representation of a neural network with two hidden layers (5 and 3 nodes, respectively). Edge thickness represents the feature weight in each node, while node size and colour represent the total outgoing node weight (normalised by outgoing node weights for the layer).

Phase 1: Small Business Loan Acceptance

Model	Grid metric	α	Training Score	AUC Test	Recall Rejected	Recall Accepted
LR	AUC	1	89.0%	71.9%	53.5%	60.2%
LR	recall macro	0.1	77.9%	71.7%	54.0%	59.9%
LR	fixed	0.001	80.0%	71.1%	55.2%	65.2%
LR	fixed	0.0001	80.1%	71.0%	55.9%	62.9%
SVM	recall macro	0.01	-	77.5%	52.6%	68.4%
SVM	AUC	10	-	89.0%	97.3%	43.3%

Small business loan acceptance results and parameters for SVM and LR grids trained on the entire dataset and tested on its "small business" subset

Model	Grid metric	α	Training Score	AUC Test	Recall Rejected	Recall Accepted
LR	AUC	0.1	88.9%	65.7%	48.5%	62.9%
LR	recall macro	0.1	78.5%	65.5%	48.6%	57.0%
SVM	recall macro	0.01	-	89.3%	47.8%	62.9%
SVM	AUC	10	-	83.6%	46.4%	76.1%

Small business loan acceptance results and parameters for SVM and LR grids trained and tested on the data's "small business" subset.

Phase 1: Small Business Loan Default

Model	Grid metric	α	Training Score	AUC Test	Recall Defaulted	Recall Paid
LR	AUC	0.1	64.8%	66.4%	65.2%	57.4%
LR	recall macro	0.01	60.4%	65.3%	64.6%	53.3%
SVM	recall macro	0.01	-	59.9%	59.8%	58.8%
SVM	AUC	0.1	-	64.2%	50.8%	65.8%

Small business loan default results and parameters for SVM and LR grids trained and tested on the data's "small business" subset.

Model	Grid metric	α	Training Score	AUC Test	Recall Defaulted	Recall Paid
LR	AUC	0.001 (L1)	69.8%	68.9%	81.0%	43.3%
LR	AUC	0.001	69.7%	69.2%	86.4%	35.0%
LR	recall macro	0.001	64.2%	69.2%	86.4%	35.0%
SVM	recall macro	0.001	-	64.1%	77.7%	48.3%
SVM	AUC	0.001	-	69.7%	77.7%	48.3%

Small business loan default results and parameters for SVM and LR grids trained on the entire dataset and tested on its "small business" subset.

Greater efficiency

- **75% default prediction**

- Given that the present loan screening has a resulting fraction of default around 20% we can infer that potentially the methodology presented in this work could

reduce the defaulting loans to 10%

with positive consequences for the efficiency of this market.

Ruled by code

In 2014, after leaving his job as FED Chairman, Ben Bernanke (who can charge $250.000 for a speech) was refused the refinancing of his home mortgage by an algorithm.

The decision was taken on the grounds that he had recently moved from a full-time employment to a self-employed status (Irwin 2014).