National Research University Higher School of Economics
Faculty of Computer Science
School of Data Analysis and Artificial Intelligence

Neural Networks in Formal Concept Analysis

Subject: Ordered Sets In Data Analysis

Presenter: M. Kirdin

Professor: Kuznetsov S.

Moscow, 2024

Outline

- Introduction
- Problem Statement
- Model Evaluation
- Results
- Conclusion
- References

Introduction

The overarching goal of this homework is to implement a **merge** of machine learning and formal concept analysis. This can be done by basing the neural network(NN) architecture on the **covering relation** (graph of the diagram) of a lattice coming from monotone Galois connections as proposed by Kuznetsov and his colleagues [1].

Problem Statement

For this task a dataset has to be chosen, its data binarized using scaling (binarization) strategy of choice, and finally the target attribute defined. Then, a **comparison** between several standard classification methods and the neural network produced as the result of this homework should be made by calculating **performance metrics** best suited for the dataset.

Model Evaluation

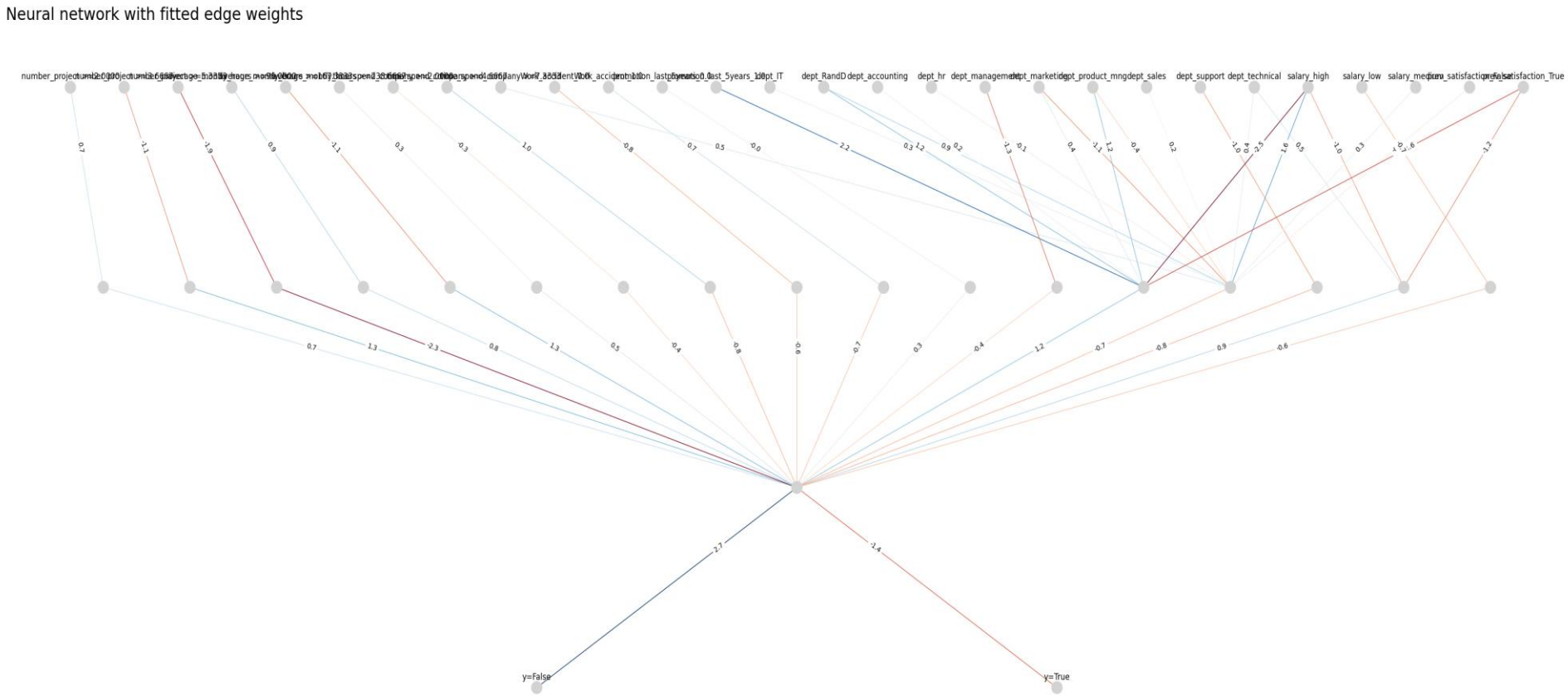
First dataset chosen was the **Employee Attrition** dataset. It describes employees' satisfaction with their workplace in a certain company.

Emp ID	satisfaction_level	last_evaluation	number_project	average_montly_hours	time_spend_company	Work_accident	promotion_last_5years	dept	salary
1	0.38	0.53	2.0	157.0	3.0	0.0	0.0	sales	low
2	0.80	0.86	5.0	262.0	6.0	0.0	0.0	sales	medium
3	0.11	0.88	7.0	272.0	4.0	0.0	0.0	sales	medium
4	0.72	0.87	5.0	223.0	5.0	0.0	0.0	sales	low
5	0.37	0.52	2.0	159.0	3.0	0.0	0.0	sales	low

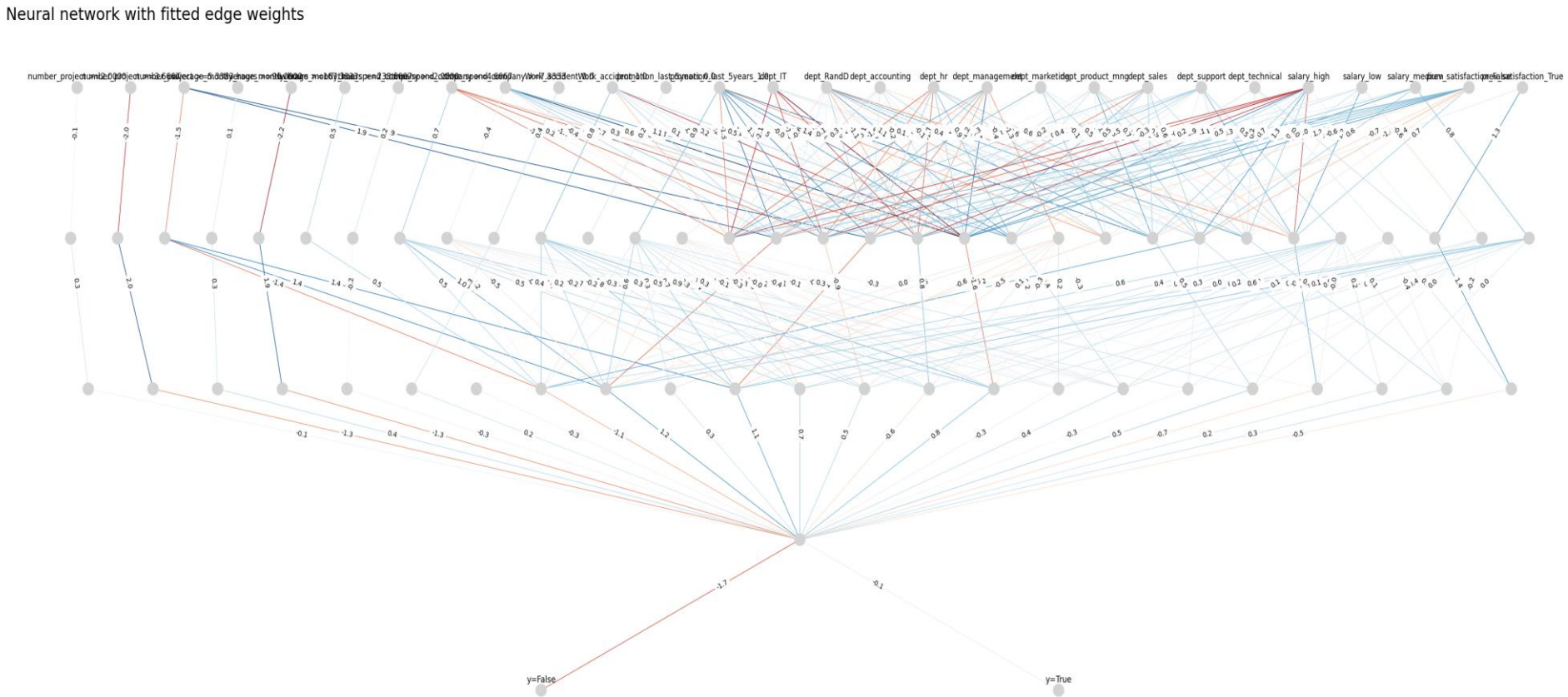
Model Evaluation

Let's see how the number of concepts used impacts the **interpretability** of a neural network

Model Evaluation

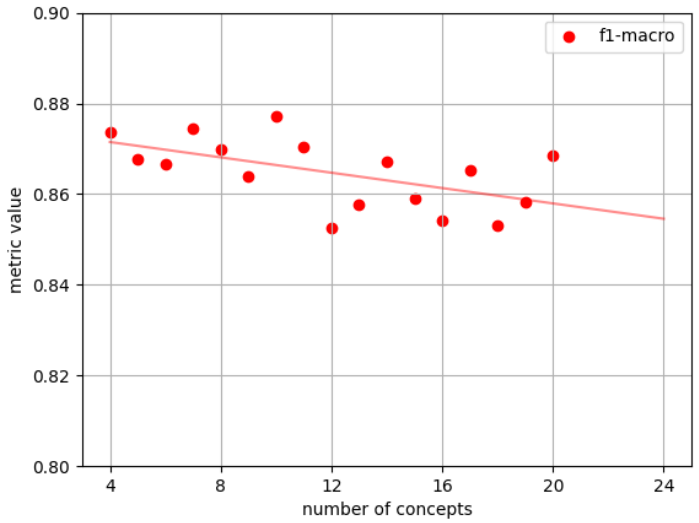


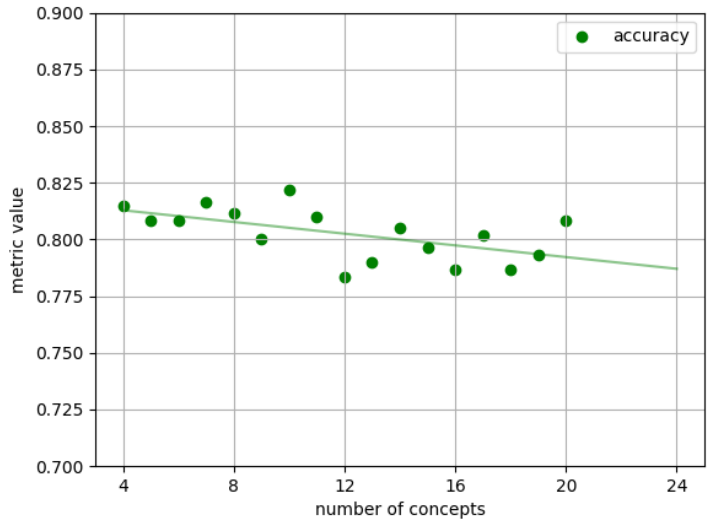
Model Evaluation



Model Evaluation

The impact of an increase in **number of concepts** used on performance metrics has to be considered too:

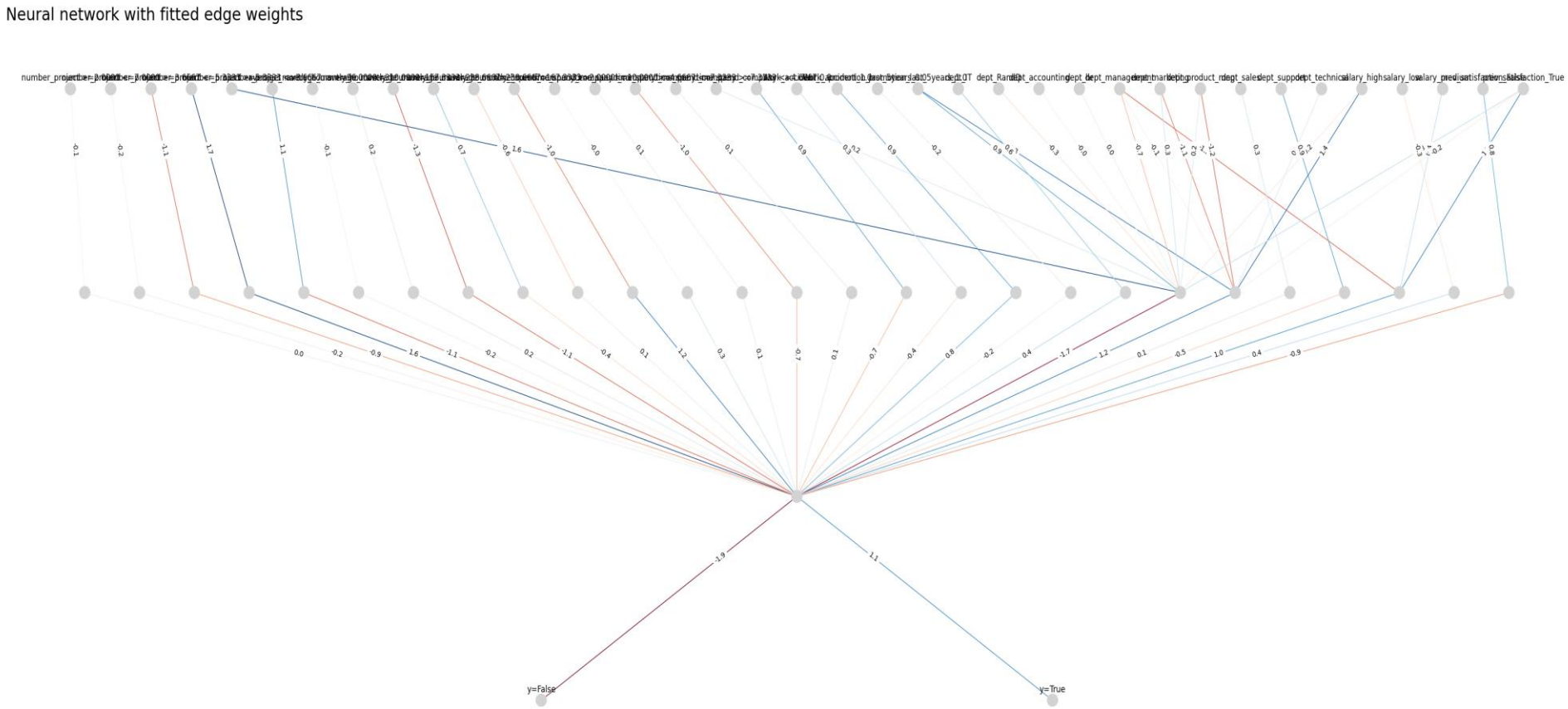




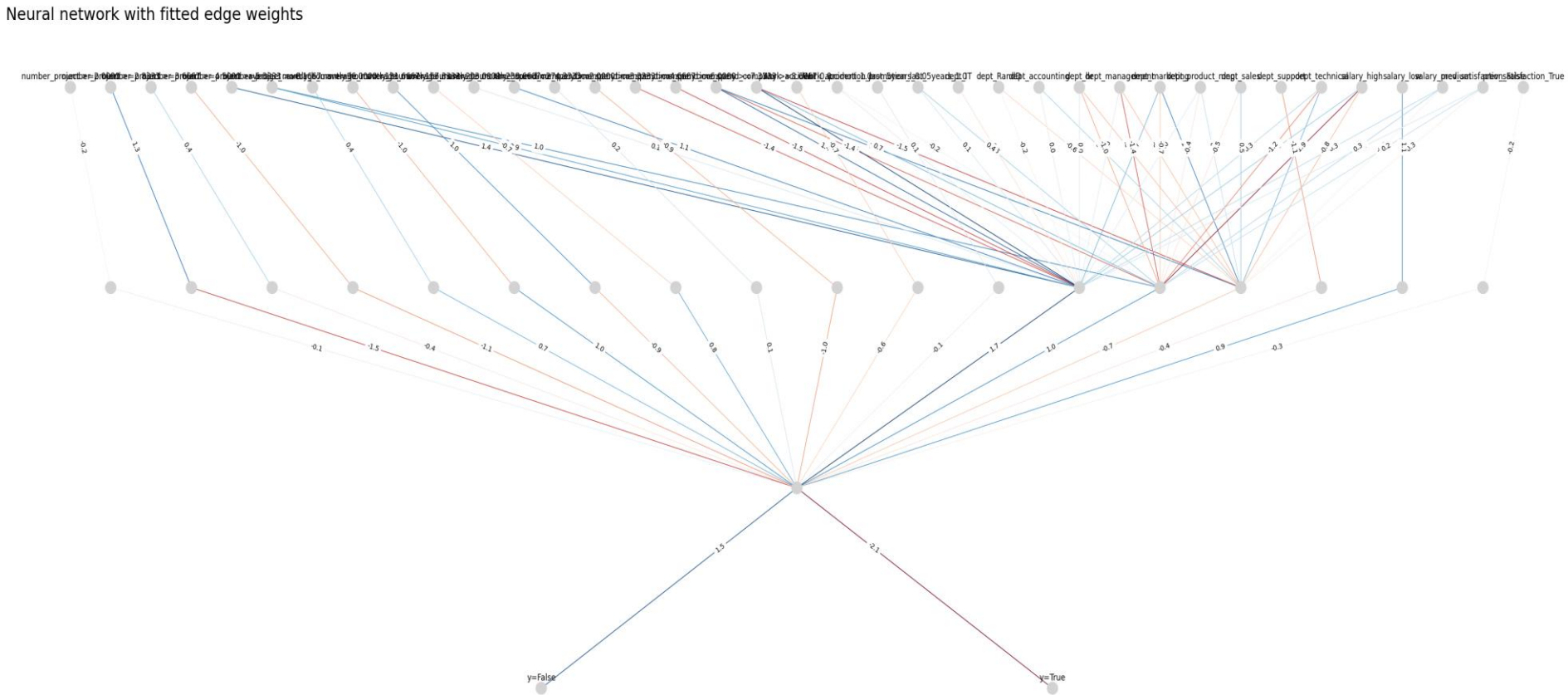
Model Evaluation

Let's also compare the impact of different **binarization strategies** on the interpretability of NN's architecture.

Model Evaluation



Model Evaluation



Model Evaluation

Let's look at the performance metrics of the neural network and several **other classifiers** on the Employee Attrition dataset.

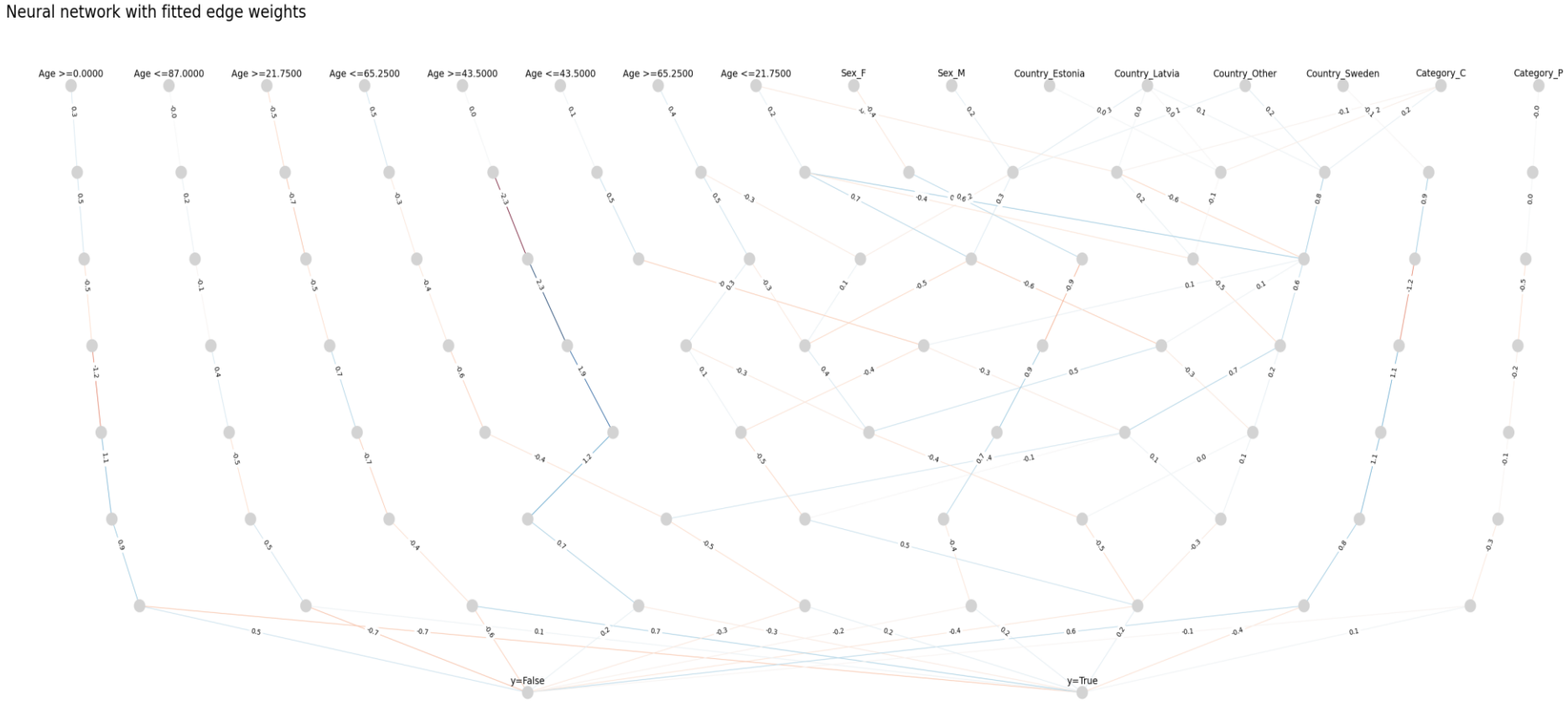
Classifier	f1-macro	accuracy
GaussianNB	0.8318	0.7567
RandomForest	0.8889	0.8383
HistGradientBoosting	0.8791	0.8300
NeuralFCA(base)	0.8643	0.8017
NeuralFCA(interordinal)	0.8707	0.8100
NeuralFCA(fine)	0.9001	0.8550

Model Evaluation

We should also check how the neural network performs on skewed data. For this task the **Estonia Disaster** dataset was used.

PassengerId	Country	Firstname	Lastname	Sex	Age	Category	Survived
1	Sweden	ARVID KALLE	AADLI	M	62	P	0
2	Estonia	LEA	AALISTE	F	22	C	0
3	Estonia	AIRI	AAVASTE	F	21	C	0
4	Sweden	JURI	AAVIK	M	53	C	0
5	Sweden	BRITTA ELISABET	AHLSTROM	F	55	P	0

Model Evaluation



Results

Finer binarization produces **better results** in terms of model metrics than both the baseline and binarization using interordinal encoding with similar number of nodes, however this comes **at the cost of interpretability**. Increasing the number of formal concepts used to construct the concept lattice does not improve the model's performance; on the contrary, it seems to correlate with decrease in classification performance. Furthermore, the NN architecture used in this work struggles with overfitting on skewed data as exemplified by the Estonia Disaster dataset.

Conclusion

Neural Networks based on concept lattices perform **on par with several ensemble models** and show promise in some machine learning applications. As exemplified by the Estonia Disaster dataset, these NNs can show connections between various features and their impact on its final decision, hence **improving human interpretability** of their inner workings.

References

1. Kuznetsov, S.O., Makhazhanov, N., Ushakov, M. On Neural Network Architecture Based on Concept Lattices // Foundations of Intelligent Systems. ISMIS 2017. Lecture Notes in Computer Science(), vol 10352, pp. 653-663. Springer, Cham. https://doi.org/10.1007/978-3-319-60438-1_64