On Neural Network Architecture based on Concept Lattices

Sergei O. Kuznetsov, Nurtas Makhazhanov, Maxim Ushakov

Moscow

16 March, 2017

Table of contents

1. General Overview
2. Generation of Formal Concepts
3. Select Concepts with best Performance
4. Construct Neural Network architecture
5. Experiments
6. Conclusions

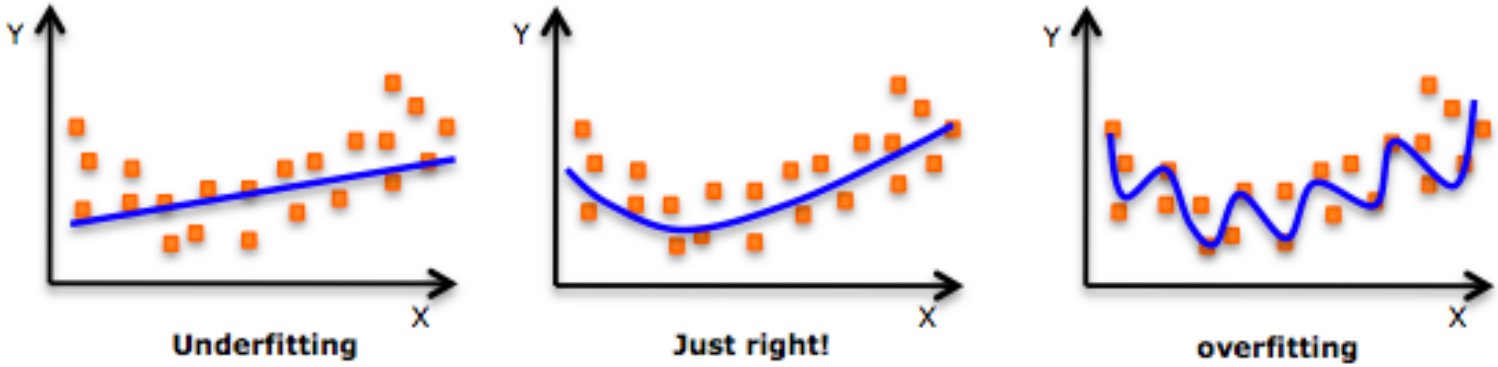
Table of contents

1. General Overview
2. Generation of Formal Concepts
3. Select Concepts with best Performance
4. Construct Neural Network architecture
5. Experiments
6. Conclusions

Motivation

Selecting an appropriate neural network architecture is a crucial problem when finding a solution based on a neural network.

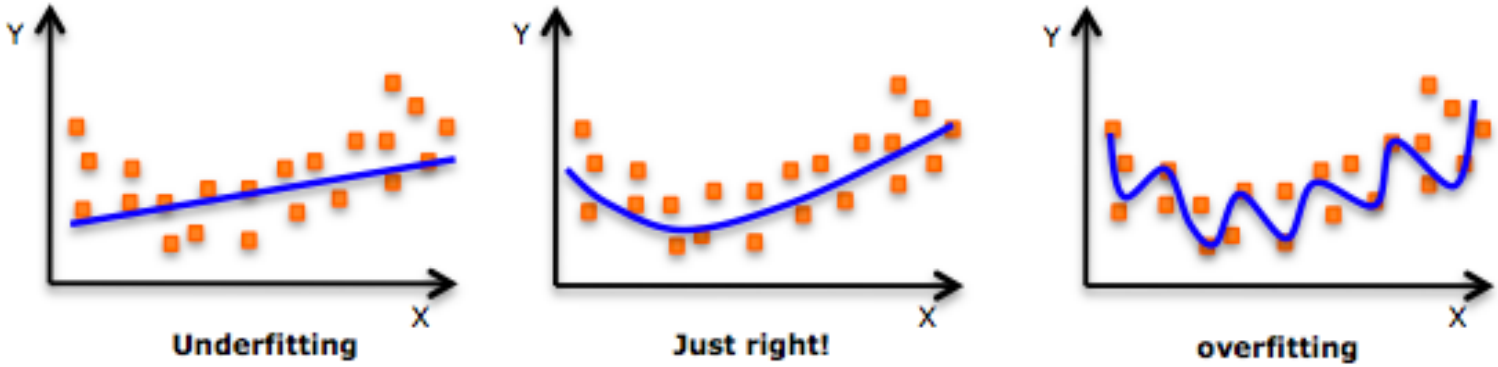
- Neural networks suffer the problem of poor interpretability of learning results.
- If the number of neurons in the network is low, then it is likely to underfit.
- If the number of neurons in the network is too high, then it is likely to overfit.



Motivation

Selecting an appropriate neural network architecture is a crucial problem when finding a solution based on a neural network.

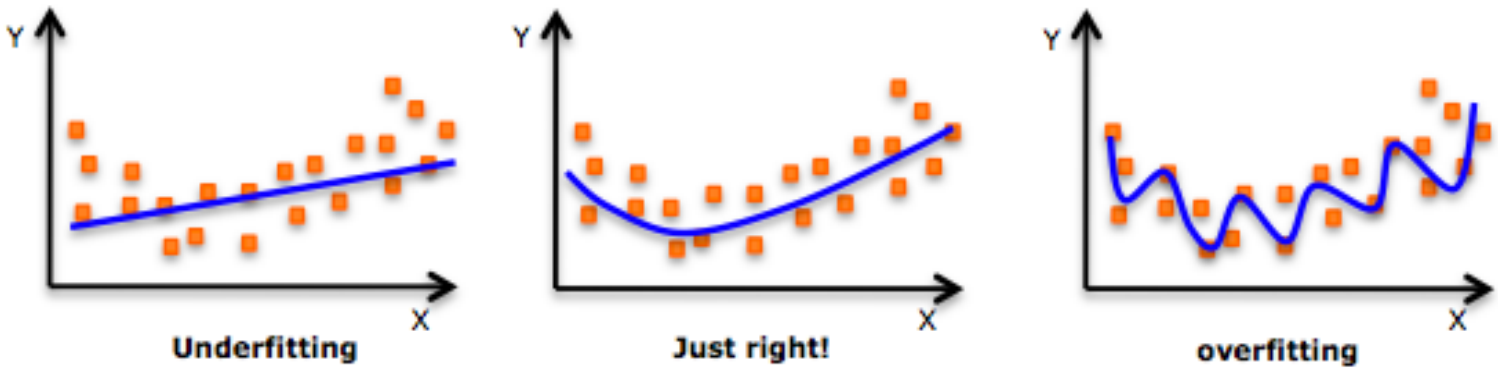
- Neural networks suffer the problem of poor interpretability of learning results.
- If the number of neurons in the network is low, then it is likely to underfit.
- If the number of neurons in the network is too high, then it is likely to overfit.



Motivation

Selecting an appropriate neural network architecture is a crucial problem when finding a solution based on a neural network.

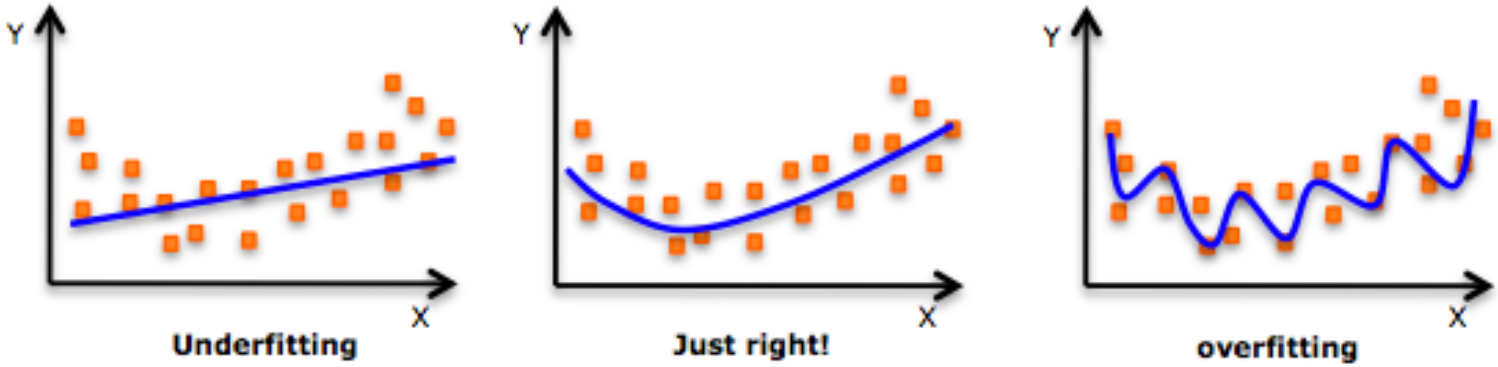
- Neural networks suffer the problem of poor interpretability of learning results.
- If the number of neurons in the network is low, then it is likely to underfit.
- If the number of neurons in the network is too high, then it is likely to overfit.



Motivation

Selecting an appropriate neural network architecture is a crucial problem when finding a solution based on a neural network.

- Neural networks suffer the problem of poor interpretability of learning results.
- If the number of neurons in the network is low, then it is likely to underfit.
- If the number of neurons in the network is too high, then it is likely to overfit.



FNN based on a Concept Lattice

We went through following steps to apply Formal Concept Analysis to constructing optimal Feedforward Neural Network architecture.

- Generate the set of formal concepts
- Select a subset of formal concepts with best performance according to some formal criterion
  - based on F-value
  - based on Score accounted on Precision and Recall
- Construct NN architecture based on selected formal concepts
- Train the obtained NN

FNN based on a Concept Lattice

We went through following steps to apply Formal Concept Analysis to constructing optimal Feedforward Neural Network architecture.

- Generate the set of formal concepts
- Select a subset of formal concepts with best performance according to some formal criterion
  - based on F-value
  - based on Score accounted on Precision and Recall
- Construct NN architecture based on selected formal concepts
- Train the obtained NN

FNN based on a Concept Lattice

We went through following steps to apply Formal Concept Analysis to constructing optimal Feedforward Neural Network architecture.

- Generate the set of formal concepts
- Select a subset of formal concepts with best performance according to some formal criterion
  - based on F-value
  - based on Score accounted on Precision and Recall
- Construct NN architecture based on selected formal concepts
- Train the obtained NN

FNN based on a Concept Lattice

We went through following steps to apply Formal Concept Analysis to constructing optimal Feedforward Neural Network architecture.

- Generate the set of formal concepts
- Select a subset of formal concepts with best performance according to some formal criterion
  - based on F-value
  - based on Score accounted on Precision and Recall
- Construct NN architecture based on selected formal concepts
- Train the obtained NN

Table of contents

1. General Overview
2. **Generation of Formal Concepts**
3. Select Concepts with best Performance
4. Construct Neural Network architecture
5. Experiments
6. Conclusions

Formal Context

The first step is generating the set of concepts. First of all, we apply scaling. Upon scaling operation each variable have become transformed to one or more binary attributes.

Context	N	G / M	color				form		firm		smooth		target
			w	y	g	b	r	$\bar{r}$	f	$\bar{f}$	s	$\bar{s}$	
	1	apple		×			×			×	×		×
K_+	2	grapefruit		×			×			×		×	×
	3	kiwi			×			×		×		×	×
	4	plum				×		×		×	×		×
	5	toy cube			×			×	×		×		
K_-	6	egg	×					×	×		×		
	7	tennis ball	×				×			×		×	
K_τ	8	mango		×			×			×	×		?

Concepts Lattice

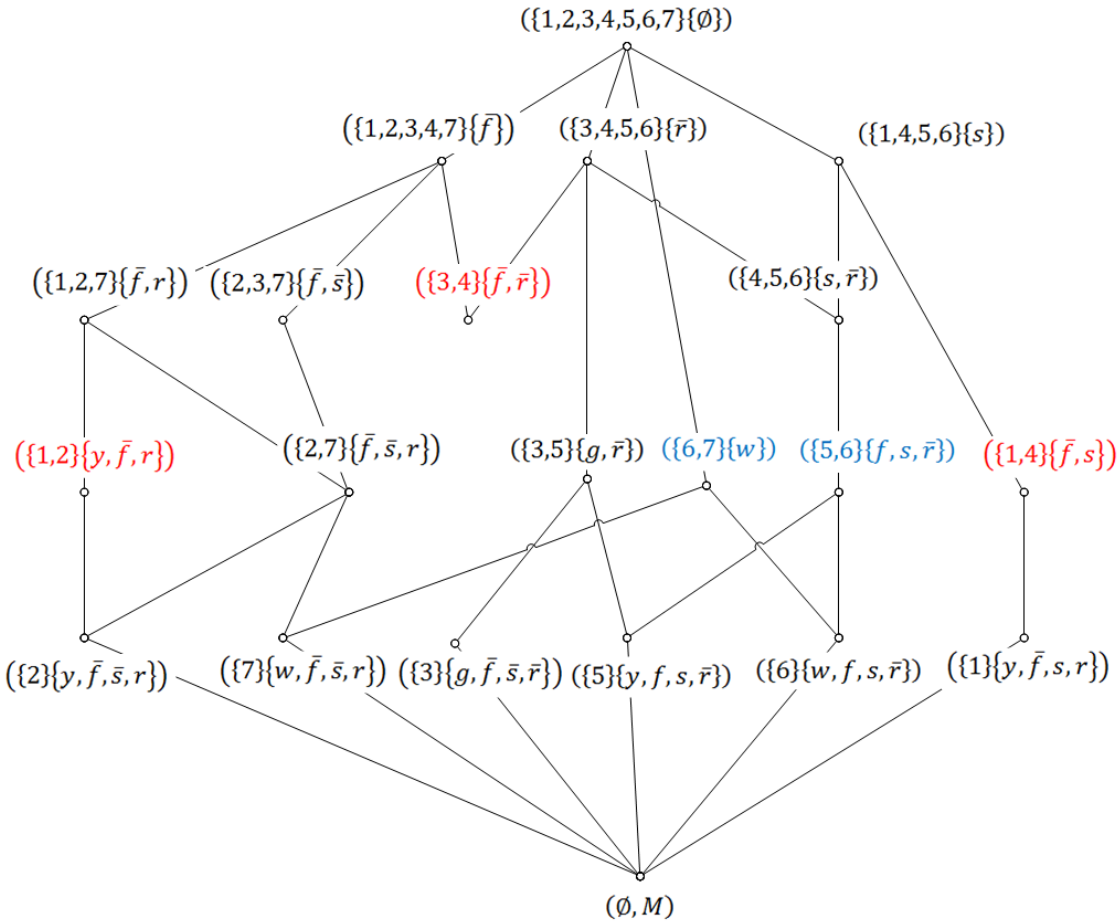


Table of contents

1. General Overview
2. Generation of Formal Concepts
3. **Select Concepts with best Performance**
4. Construct Neural Network architecture
5. Experiments
6. Conclusions

Select Concepts with best Performance

Upon generating formal concepts, it is necessary to calculate performance of each concept on the data set and select the best. Here we consider two types of calculating performance:

- based on F-value $F(h_i) = 2 * \frac{Precision(h_i) * Recall(h_i)}{Precision(h_i) + Recall(h_i)}$
- based on Score accounted on Precision and Recall
  $Score(h_i) = \alpha Precision(h_i) + (1 - \alpha) Recall(h_i)$

Then we will select the best ones. Here, we will consider two techniques for selecting the best concepts:

- based on Score threshold
- based on concept "usefullness" to cover undetermined examples

Select Concepts with best Performance

Upon generating formal concepts, it is necessary to calculate performance of each concept on the data set and select the best. Here we consider two types of calculating performance:

- based on F-value $F(h_i) = 2 * \frac{Precision(h_i) * Recall(h_i)}{Precision(h_i) + Recall(h_i)}$
- based on Score accounted on Precision and Recall
  $Score(h_i) = \alpha Precision(h_i) + (1 - \alpha) Recall(h_i)$

Then we will select the best ones. Here, we will consider two techniques for selecting the best concepts:

- based on Score threshold
- based on concept "usefullness" to cover undetermined examples

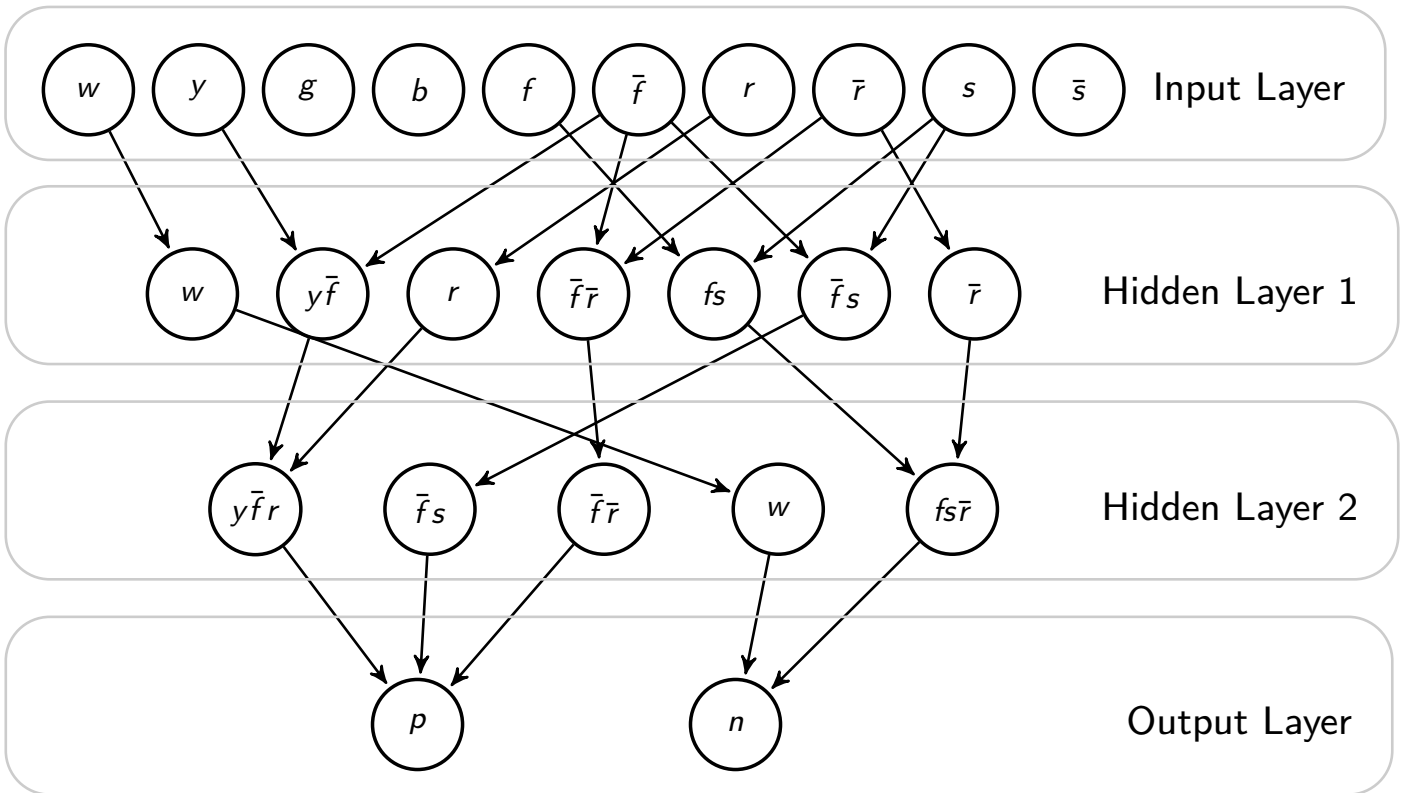
Table of contents

1. General Overview
2. Generation of Formal Concepts
3. Select Concepts with best Performance
4. **Construct Neural Network architecture**
5. Experiments
6. Conclusions

Neural network architecture

After selecting best concepts we construct the neural network from them using the covering relation of the upper part of the lattice diagram. The network would consist of the following layers:

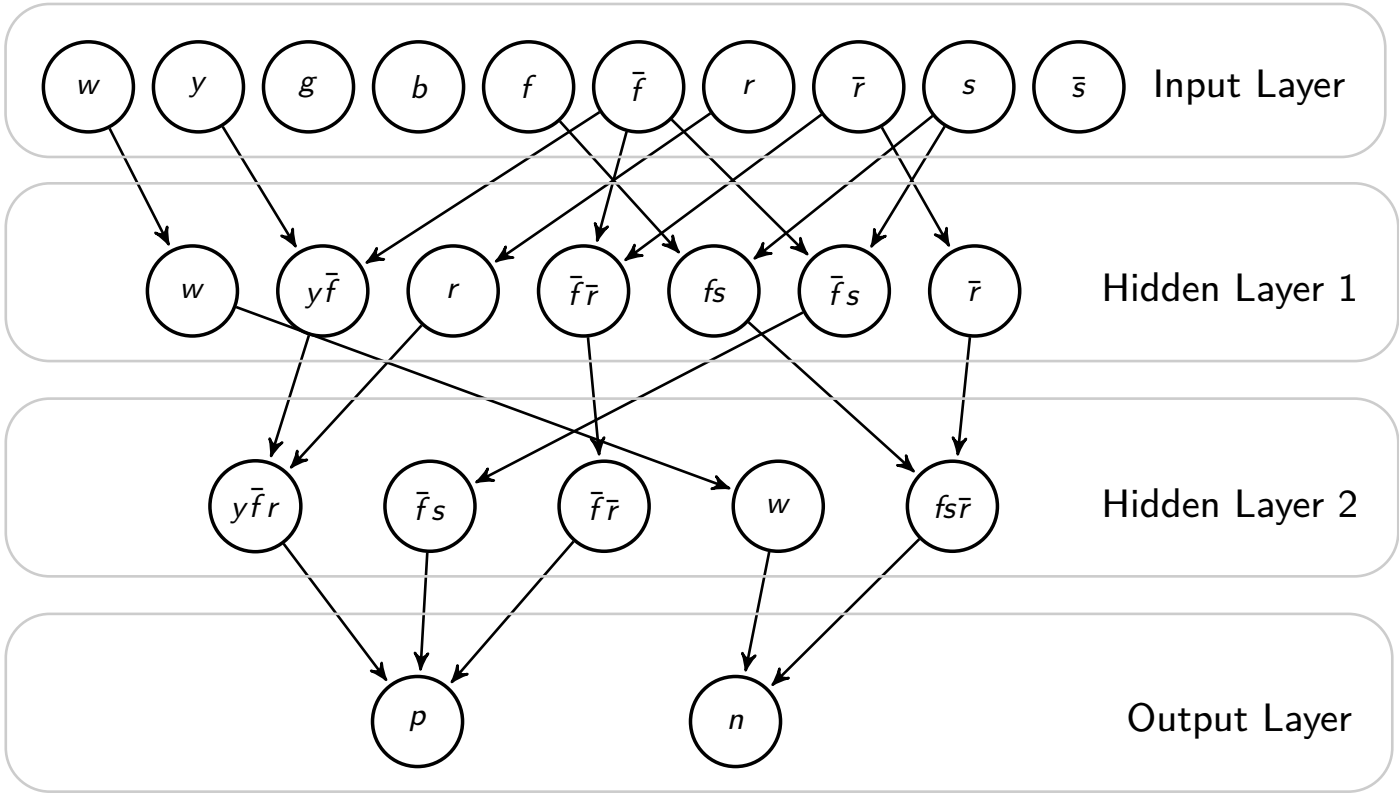
- Input layer *Inp* consists of neurons that are related to attributes $m \in M$ of context $K = (G, M, I)$.



Neural network architecture

After selecting best concepts we construct the neural network from them using the covering relation of the upper part of the lattice diagram. The network would consist of the following layers:

- Hidden layers Hid_i. Each concept can be uniquely represented by its intent. We iteratively connect attributes from M in hidden layers so that on the *last hidden layer* we can achieve neurons corresponding to previously selected concepts.



Neural network architecture

After selecting best concepts we construct the neural network from them using the covering relation of the upper part of the lattice diagram. The network would consist of the following layers:

- Output layer *Out*. Output layer will be of a softmax type. In the case of binary classification, there will be only two neurons: neurons that predict positive and negative classes.

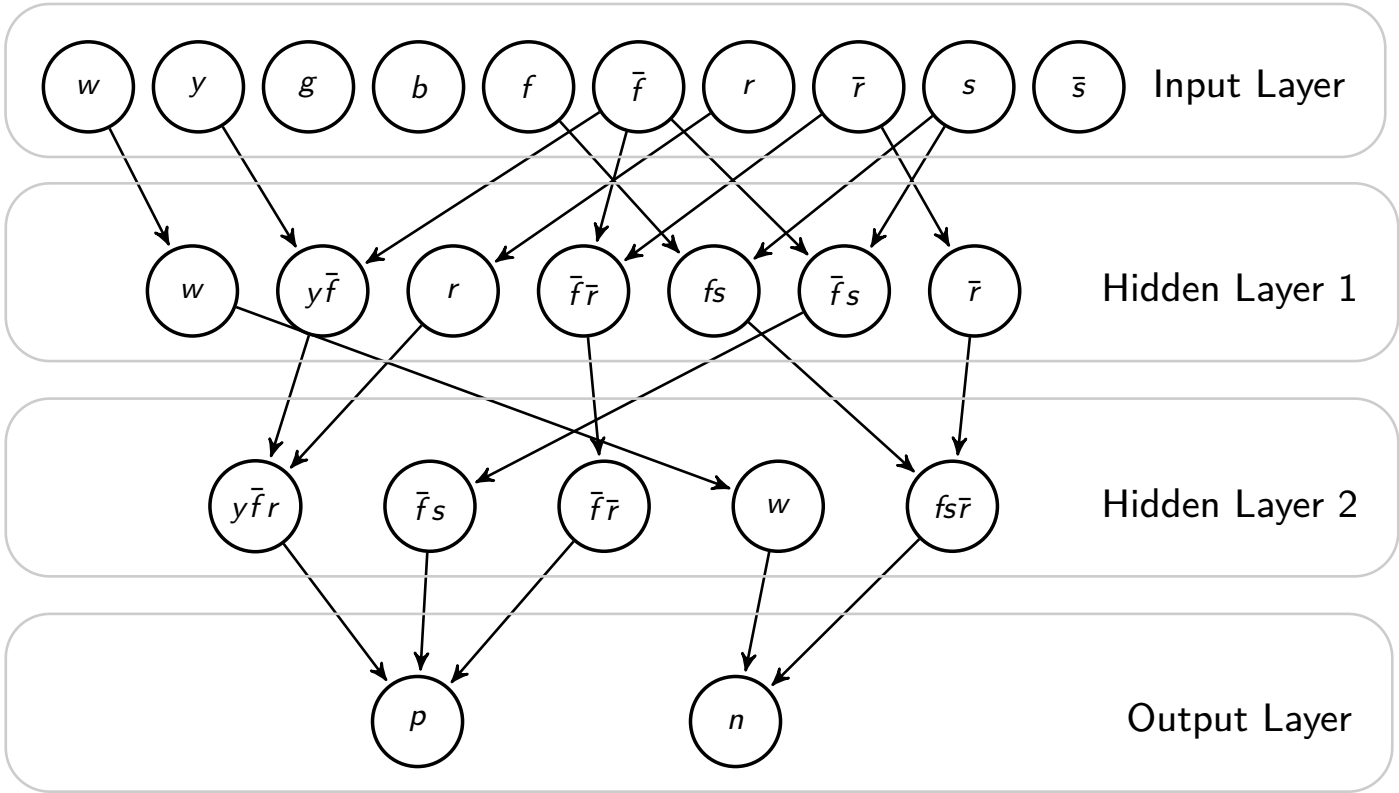


Table of contents

1. General Overview
2. Generation of Formal Concepts
3. Select Concepts with best Performance
4. Construct Neural Network architecture
5. **Experiments**
6. Conclusions

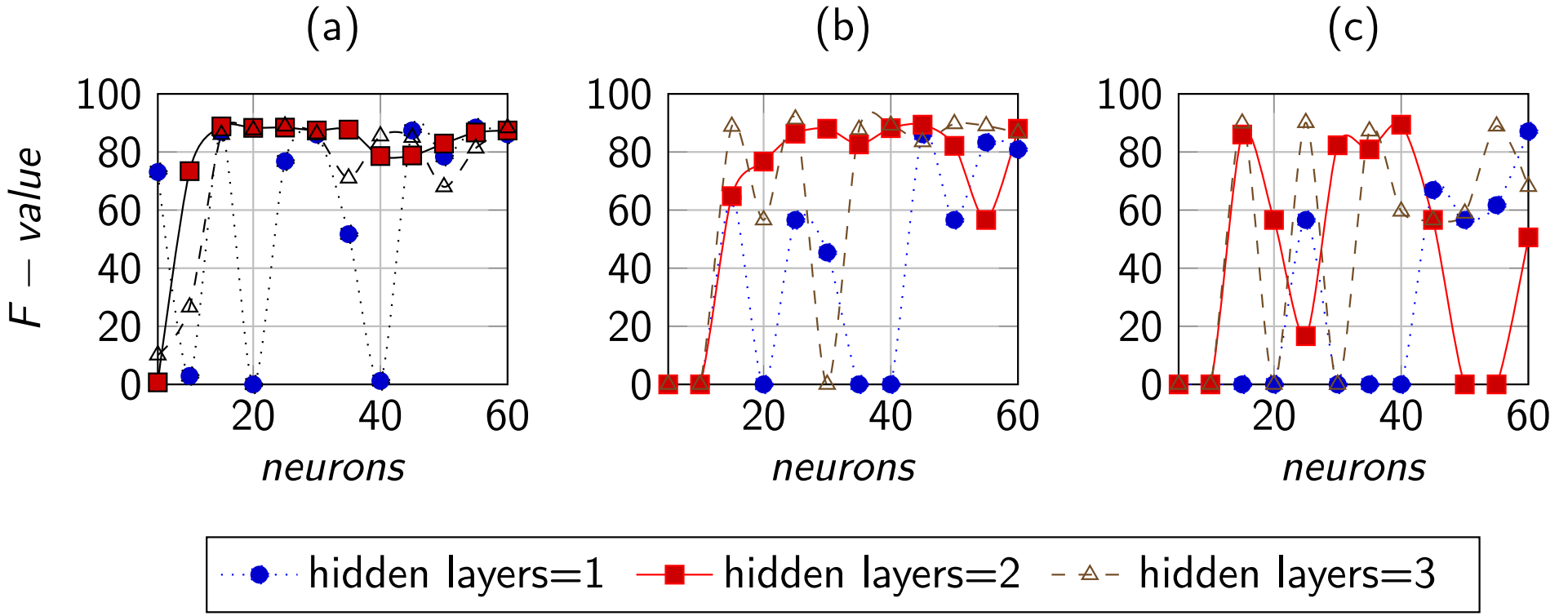
Data

We have taken the data from the open source UCI Machine Learning Repository (http://archive.ics.uci.edu/ml/index.html) for experiments. We selected six datasets:

Dataset	Train sample	Test sample	Number of variables	Number of classes in target variable
Breast Cancer	512	57	30	2
CreditCard Default	27000	3000	23	2
Heart Disease	273	31	13	5
Mammographic Mass Data	865	96	5	2
Seismic Bumps	2326	258	18	2
Car Evaluation	1555	173	6	4

Experiments

Performance of NNs with different architectures on breast cancer data. On x-scale number of neurons in each hidden layer, on y-scale F-values. Each NN were learned applying Adam stochastic optimization strategies with initial learning rates equal: (a) 0.01, (b) 0.001, (c) 0.0001.



Experiments

Different Machine Learning methods Performance.

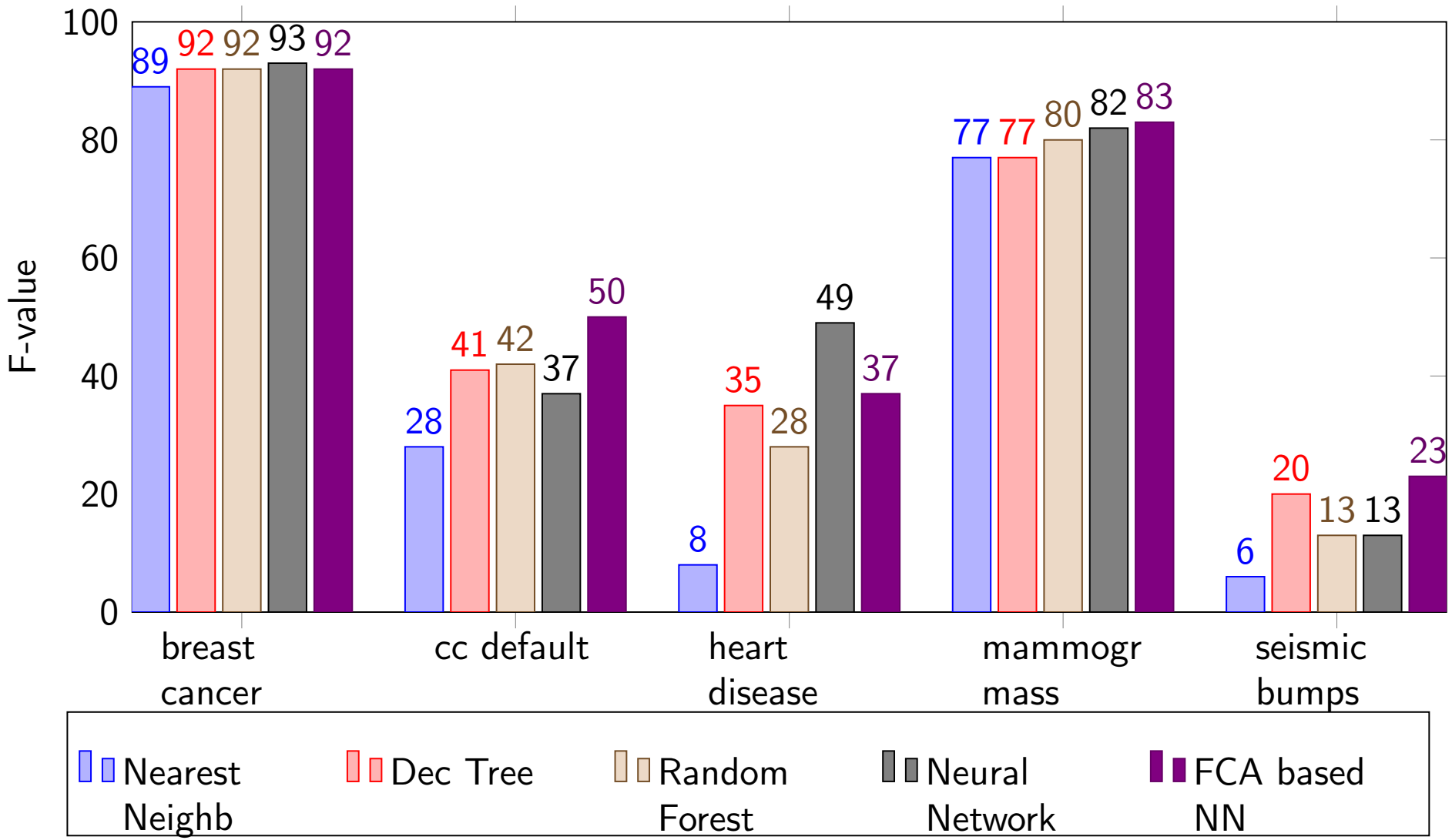


Table of contents

1. General Overview
2. Generation of Formal Concepts
3. Select Concepts with best Performance
4. Construct Neural Network architecture
5. Experiments
6. **Conclusions**

Conclusions

- Neural networks that are based on concept lattices are very sparse comparing with standard fully-connected networks. All neurons in the last hidden layer are related to concepts coming from the dataset.
- Another advantage of neural networks based on concept lattices is their interpretability, which is very significant in domains like credit scoring. One can both predict the probability of default of applicants but also implement specific rules and then weigh them according to their importance for predicting target variable. Thus, neural networks based on concept lattices can be implemented in domains where it is important to explain why some or other object assigned as that particular class.

Thank you!