BABEŞ-BOLYAI UNIVERSITY
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Thesis Title

B.Sc. Thesis

PhD student: name
Scientific supervisor: Lect. Ph.D. Radu D. Găceanu

2015

Abstract

Abstract here...

Contents

List of Figures

List of Tables

Introduction

Introduction goes here...

Chapter 1

Theoretical background

Aici un rezumat al capitolului. Este descrisa fiecare sectiune. In Section 1.1 blabla.

1.1 Section 1

A neural network classifier may be considered as a mapping

$$F : R^d \rightarrow R^M, \tag{1.1}$$

where x represents the input of the network and the output y is the classification result.

Pentru a folosi modul matematic - cu beginalign sau intre dous semne dolar.

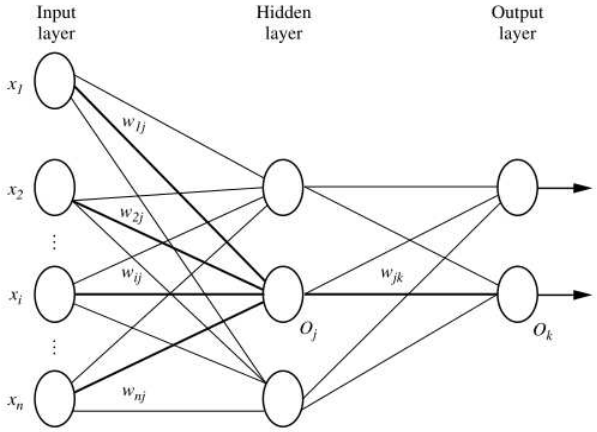

Figure 1.1: A two-layer feed-forward neural network. Image source: [2].

A multilayer feed-forward neural network is shown in Figure 1.1.

In caption-ul figurii la image source trebuie citata cartea [2] sau articolul de unde este luata figura sau se scrie link-ul catre figura.

A good clustering algorithm should satisfy several requirements as follows [4]:

- scalability, i.e., it should still work properly when applied to large datasets
- ability to handle different types of features, namely, continuous, binary, categorical, ordinal and ratio-scaled
- tolerance to high dimensional data
- independence with respect to the cluster shape

- minimal domain knowledge requirements, i.e., not too many parameters that need to be initialized
- tolerance to noise
- tolerance to outliers
- insensitivity to the order in which data items are read
- interoperability
- usability.

Definition 1.1.1 *The expected information needed for classifying a data item in D (the entropy of D) is given by*

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i), \tag{1.2}$$

where:

- p_i *denotes the probability that a data item from the dataset belongs to the class* C_i
- $p_i \approx \frac{|C_i|}{|D|}$.

According to Definition 1.1.1 blabla.

Remark 1.1.1 *The goal is to find the attribute A from which* $Info_A(D)$ *is minimal and hence the purity of the obtained partitions is maximal.*

According to Remark 1.1.1 bla.

Bayesian classification is a statistical classification method which computes the probability that a given data item belongs to a particular class thus predicting class memberships [1].

Theorem 1 (Bayes' theorem) . *Let us denote by* $C_j, j = \overline{1, J}$ *the possible classes and by* $P(C_j \mid X_1, X_2, \ldots, X_p)$ *the posterior probability of belonging to the class* C_j *given the features* $X_1, X_2, \ldots, X_p$ *then*

$$P(C_j \mid X_1, X_2, \ldots, X_p) = \frac{P(X_1, X_2, \ldots, X_p \mid C_j) \cdot P(C_j)}{\sum_j P(X_1, X_2, \ldots, X_p \mid C_j)},$$

where:

- $P(X_1, X_2, \ldots, X_p \mid C_j)$ *denotes the probability of an item with individual characteristics (features)* $X_1, X_2, \ldots, X_p$ *belonging to the class* C_j
- $P(C_j)$ *denotes the unconditional prior probability of belonging to the class* C_j.

The goal of the rule based classifier is to construct the smallest set of rules such that consistency with respect to training data is preserved. A large number of rules is an indication of attempting to remember the training set, as opposed to discovering the assumptions that govern it [3]. A general pseudo-code for rule-based classifiers is presented in Algorithm 1.

Algorithm 1 Learn rules

1: $RuleSet \leftarrow \emptyset$

```
2: for all classes c do
3:    repeat
3:       Rule ← Find_Best_Rule(c)
3:       Remove items covered by Rule
4:    until termination_condition
5:    RuleSet ← RuleSet ∪ {Rule}
6: end for
end
```

MisclassificationId	Similarity	C1	C2	C3
106	0.25	0	0.95	0.84
119	0.24	0	0.96	0.83
133	0.19	0	1	0.89
134	0.24	0	0.96	0.82

Table 1.1: Cluster2 — RepresentativeId (95)

From the table 1.1 blabla.

Chapter 2

Conclusions

Conclusions here...

Bibliography

[1] Irad Ben-Gal. *Bayesian Networks*. John Wiley & Sons, Ltd, 2008.

[2] Jiawei Han and Micheline Kamber. *Data Mining: Concepts and Techniques, 2nd ed.* Morgan Kaufmann, 2006.

[3] Sotiris B. Kotsiantis. Supervised machine learning: A review of classification techniques. *Informatica (Slovenia)*, 31(3):249–268, 2007.

[4] Stephen Marsland. *Machine Learning: An Algorithmic Perspective*. Chapman & Hall/CRC, 1st edition, 2009.