

TransDeep: Transparent Deep ensemble method based on argumentation for classification

Naziha Sendi^{*,†}, Farida Zehraoui^{*}, Nadia Abchiche-Mimouni^{*}, Etienne Colle & François Teboul[†]
^{*}IBISC, Univ Evry, Université Paris-Saclay
[†]BeWellConnect

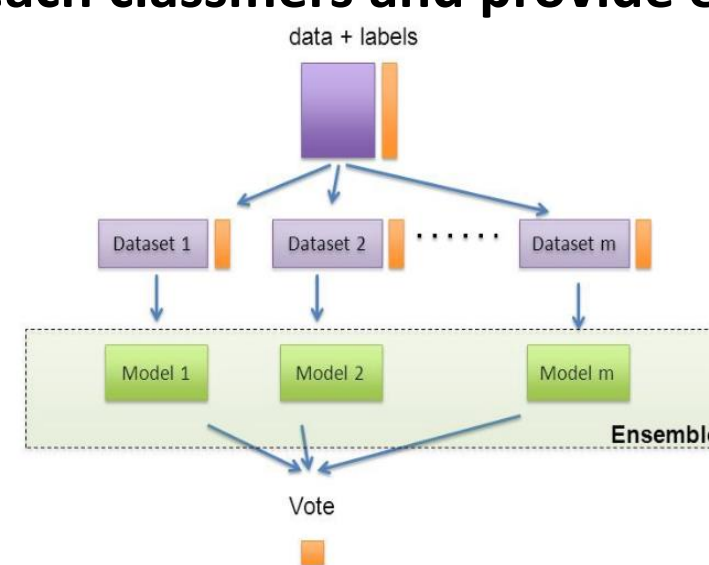
Motivation

The goal of any machine learning problem is to find a single model that will best prediction. Rather than using one model and hoping this model will be the best/most accurate predictor, ensemble methods improve machine learning results by combining different models.

One of the major criticisms of machine learning algorithms is their being inexplicable, since they do not provide results explanation.

With the development of the machine learning, the explanation behind every classification has come a necessity and nowadays the goal is to understand the prediction results.

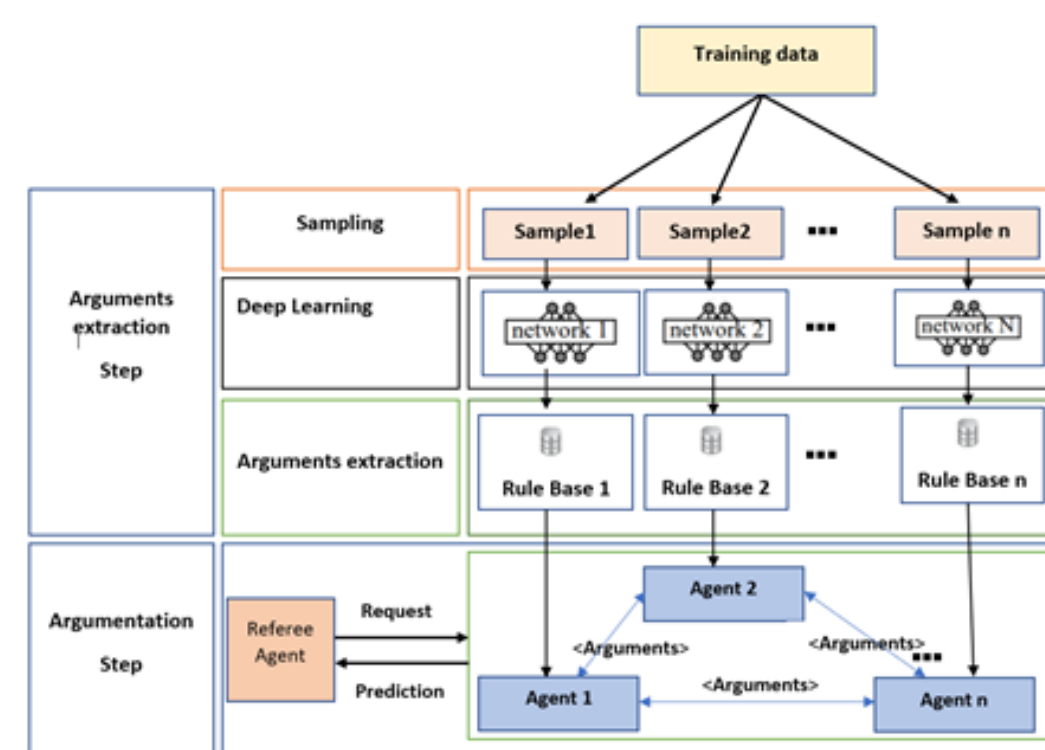
This work presents *TransDeep*, a transparent deep ensemble method based on argumentation that combines machine learning algorithms with multiagent system to improve classification and to explain the predictions. The idea is to extract automatically arguments from classifiers and combine them using argumentation in order to exploit the internal knowledge of each classifiers and provide explanation behind the decisions.



- ✓ Ensemble method integrate only classification results rather than internal classification knowledge.
- ✓ No explanation behind decision is given,

Method

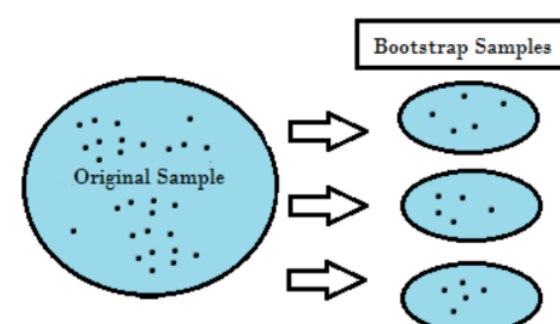
TransDeep is an original approach for explaining predictions. It is based on extracting knowledge from classifiers and then combining them by using multi agent's argumentation.



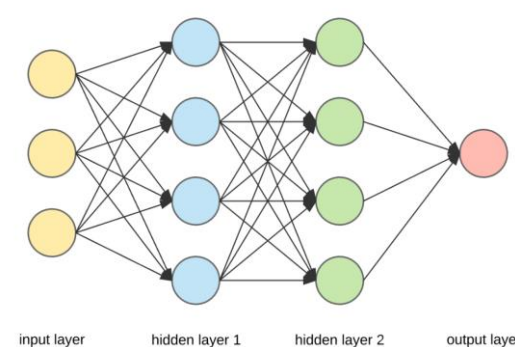
Two steps:

1. Arguments extraction step

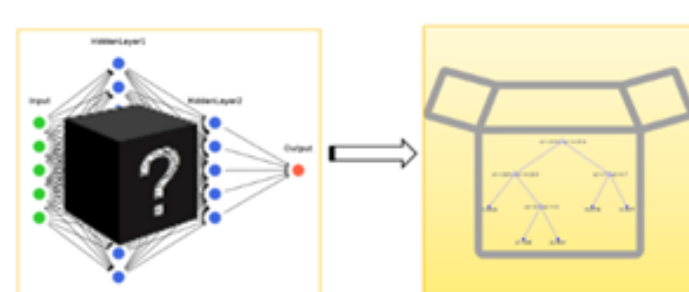
- ❖ **Bootstrap Sampling:** Several samples of the same size are generated repeatedly with replacement from the training data.



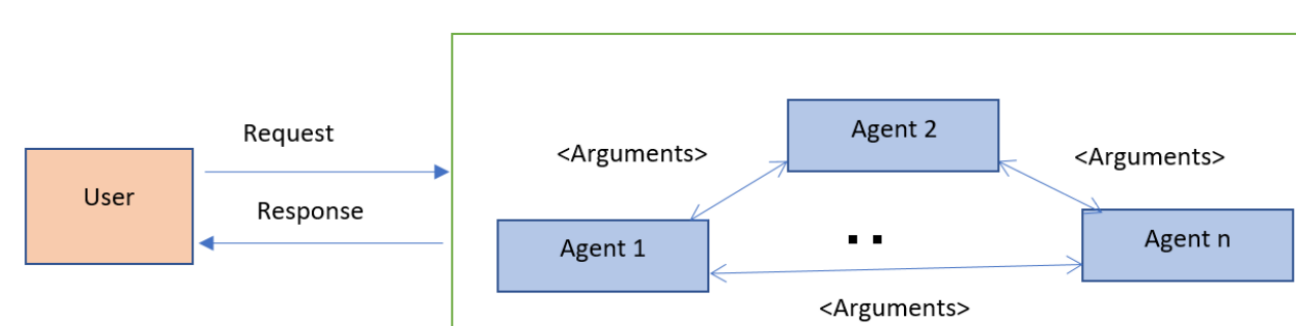
- ❖ **Deep Learning:** Each sample is used by a deep neural network, which is one of the significant breakthrough and powerful techniques in data mining and machine learning algorithms.



- ❖ **Rules extraction from Neural networks:** Using a rule extraction algorithm in order to characterize and interpret them. A rule base is generated from each deep neural network,



2. Argumentation step: Each rule base is integrated into an agent. Agents use individual knowledge, negotiate by exchanging their arguments to provide a prediction.



Results

A specific dataset that is a realistic virtual population is used:

- ✓ with the same age, sex and cardiovascular risk factors profile than the French population aged between 35 and 64 years old.
- ✓ a temporal list of visits is associated to everyone.
- ✓ 40000 individuals monitored for hypertension during 10 visits per individual.
- ✓ Each visit contains: the systolic blood pressure (SBP), diastolic blood pressure (DBP), etc.
- ✓ The goal is to predict the optimal treatment.

- Improvement of predictions

Ensemble methods	Accuracy
Adaboost	0.70
Bagging	0.72
DIMLP-B	0.79
DIMLP-A	0.85
Our approach Without expert knowledge	0.83
Our approach With expert knowledge	0.89

- Explanation of the raison behind the predictions: not only a prediction is given, but also the reasons behind that classification are provided to the user as well in a form that is easy to grasp.

The treatment for the patient x is class= BB

R1: (age<50)(SBP3>137.027)(DBP3>83.85)->Class=BB(CD=387)

- Possibility of injecting recommendations that are given by a medical expert: This is very crucial for a decision-making system that is why, injected medical recommendations for hypertension treatment is provided by *TransDeep*.

Examples of medical recommendations:

- ❑ If age < 50 year → the doctor chooses between BB, IEC and SAR
- ❑ If age > 50 year → the doctor chooses between DI and AC

Conclusion

TransDeep

- Is a new deep ensemble method based on argumentation.
- Integrates internal classification knowledge in addition to the classification results.
- Improves the accuracy of the predictions.
- Allows the user to judge whether a result is acceptable or not.
- Offers the possibility of injecting expert domain knowledge.