Conflict Resolution through Machine Learning

Veera Venkata Subramanya Aditya Bhamidipati
Naveen Kumar Vadlakonda
Sneha Gavini

Department of Computer Science
Georgia State University
Atlanta, Georgia 30303
vbhamidipati1@student.gsu.edu
nvadlakonda1@ student.gsu.edu
sgavini1@ student.gsu.edu

Abstract— **Decision intelligence is a new academic attraction concerned with all aspects of selecting between options. Making ever-faster decisions leveraging a vortex of data in ecosystems that are in constant motion, requires an assembly of increasingly convoluted techniques. Such an approach was shown to weaken the maintenance, scalability, and flexibility of both processes and decisions. Because negotiation and conflict resolution are Convoluted and unstructured tasks, they need sophisticated decision support. A crucial characteristic of such support is systems that can improve their efficiency of solution quality by employing machine learning techniques — the intent of this paper towards the application of conflict resolution through machine learning.**

Keywords— Decision Intelligence, Conflict Resolution, Machine Learning, Artificial Intelligence, Naïve Bayesian, Decision Trees, Random-Forest, Perceptron Training

I. INTRODUCTION

Decision intelligence is an engineering discipline augmenting data science from social science, decision theory, and managerial science. The application of which provides a framework for best practices in organizational decision-making and processes for applying machine learning at scale. The basic idea is that decisions based on our understanding of how actions lead to outputs. Decision intelligence is a discipline for analyzing a link of cause and effect, and decision modeling is a visual language for representing these links. Though many elements of decision intelligence, such as sensitivity analysis and analytics, are mature disciplines, they are not in extensive use by decision-makers.
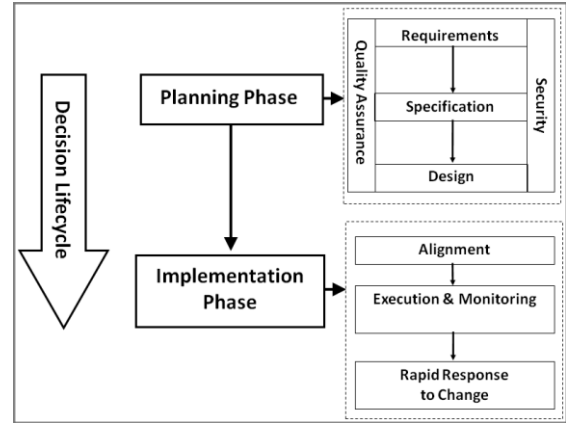

Fig.1 Framework for Decision Intelligence

Decision intelligence is a "multi-link" extension to artificial intelligence, which is in full use for single-link analysis. From this point of view, machine learning appears as answering the question, "If I know/see/hear X, what can I conclude?" whereas decision intelligence answers: "If I take action X, what is the outcome?". The latter question usually involves chains of events, sometimes including complex dynamics like feedback loops. In this way, decision intelligence unifies complex systems, machine learning, and decision analysis.

II. BACKGROUND

A key area where AI and machine learning can create value in companies today is the acceleration of the decision-making process. An affordable data storage and higher computational processing is giving machine learning the power to analyze vast data sets in the direction that delivers more accurate results. Due to the size and convolution of these data sets, machine learning can help untangle value from this data in a manner that humans cannot. As a result, machine learning is presently able to guide better business decisions and more intelligent courses of action with minimal human involvement.

III PROJECT DESCRIPTION

Decision Intelligence is applied to the "Bank Loan" dataset using various machine learning models, and extensive research is conducted on variation in the model behavior with the diversity on dataset engineering. The Model Experiments are as follows:

Model I - Naïve Bayesian Classifier

Naïve Bayesian classifier applies Bayes' theorem with the naive assumption of independence between every pair of features. Naive Bayes is a group of probabilistic algorithms that take advantage of probability theory and Bayes' Theorem to predict the tag of a text. Naïve Bayesian classifier assumes that the state of an attribute depends on the decision output. The model is applied on a "Bank Loan" dataset to decide the sanction of Personal loan and the performance trends of model concerning change in size of dataset, number of features, and variation in test splits shown in Tab.2, Tab.3 and Tab.4.
Observations from testing the model are:
1. As we decrease the size of training data and increase the size of the test data the accuracy decreases.
2. Accuracy values remain constant when features like age, zip code, family removed from dataset.
3. The accuracy is in direct proportion with the size of the dataset.

Model II - Decision Tree Classifier

Decision tree is a decision support tool using a tree-like graph and their possible outcomes, including probable event outcomes, resource values, and utility. It is one way to project an algorithm that contains conditional control statements. The depth from root to leaf represent classification rules. It splits a dataset into smaller and smaller subsets while at the same time, an associated decision tree is incrementally developed. The algorithm uses *Entropy* and *Gini Index* measures to construct a decision tree.

Entropy is the uncertainty of a random variable. It characterizes the impurity of a variable collection of examples.
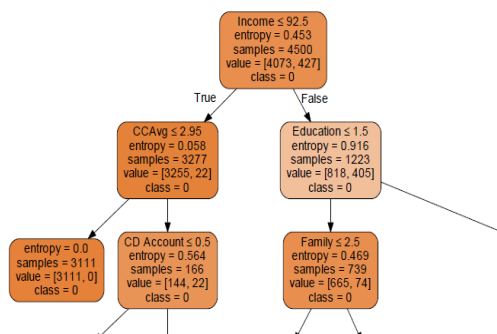
$$Entropy = \sum_{i=1}^{C} -p_i * \log_2(p_i)$$

Fig.2 Classification of decision tree based on Entropy

Gini Index measures how often a randomly selected element being incorrectly identified. The Gini Index is calculated by differencing the sum of the squared probabilities of each class from one.
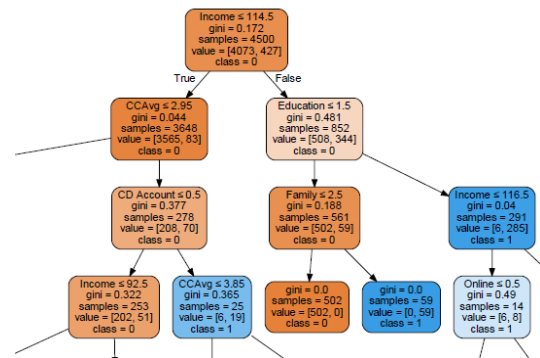
$$Gini = 1 - \sum_{i=1}^{C} (p_i)^2$$

Fig.3 Classification of decision tree based on Gini Index

The trends in accuracy basing on Gini Index and Entropy observations recorded in Tab.2, 3, 4

Observations from testing the model are:
1. The higher the entropy, the more is the crudeness of data, and hence less is the accuracy of the model.
2. Gini Index favors large partitions, which means that the dataset sample consideration should be high.

Model III - Random Forest Classifier

Random forests are bagged decision tree models that split on a subset of features on each split. Random forests is an ensemble method where several decision tree algorithms are built with different subsets of the training data and then find the most repeated or standard answer to an instance. The random forest algorithm randomly selects observations and features to build several decision trees and then averages the results. The algorithm also utilizes bootstrap aggregating, also known as bagging, to reduce overfitting and improve generalization accuracy Bagging refers to fitting each tree on a bootstrap sample rather than on the original sample.

Use Cases of Random Forests:

1. Random Forest is suited for multiclass problems.
2. Random Forests Works well with a mixture of numerical and categorical features.
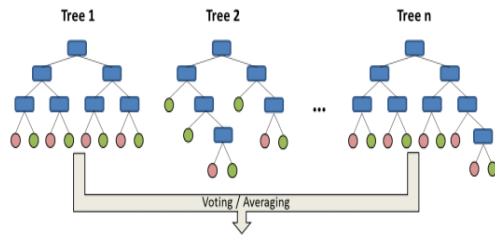
Fig.4 Random Forest with two trees

Model IV – Simple Neural Network - Supervised

The perceptron was architected for a specific type of Machine Learning problems: binary classification problems. The reason why it solves solely binary classification problems is because of the nature of its activation function. The perceptron consists of 4 parts.

- Input values or One input layer
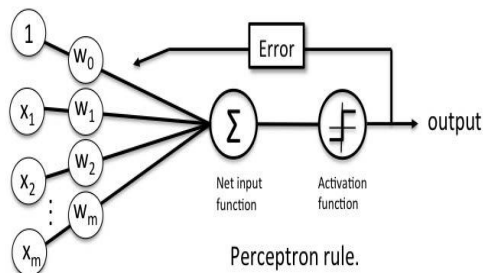- Weights and Bias
- Net sum
- Activation Function

Fig.5 Simple Perceptron Modeling

The functioning of perceptron is mainly based on the number of epochs the network is run for and the trends based on the number of epochs is discussed in Tab.1

S. No	Model	500	50	1
1	Perceptron	93.447	91.437	90.076

Tab.1 Perceptron Behaviour based on number of epochs

Use Cases of Perceptron:
1. Classify data by recognizing patterns.
2. Detect anomalies or novelties, when test data does not match the usual patterns.
3. To Approximate a target function–useful for predictions and forecasting.

IV REPORT ON MODEL ENGINEERING

Exhaustive engineering has been applied to the four machine learning models discussed above, and their trends based on the size of dataset, number of features, and Split ratio of the test set are recorded.

On Comparing the accuracies of various machine learning models with various experimentations, the results are as follows:

S. No	Model	5000 Records	500 Records
1	Naïve Bayes	89	74
2	Decision Tree	Gini - 98.8	Gini – 98.0
		Entropy – 98.2	Entropy–98.0
3	Random Forests	Gini - 99.0	Gini - 96.0
		Entropy - 99.4	Entropy - 94.0
4	Perceptron	93.447	89.357

Tab.2 Comparison of Accuracy based on dataset size

S. No	Model	12 Features	9 Features
1	Naïve Bayes	89	89
2	Decision Tree	Gini - 98.8	Gini – 95.4
		Entropy – 98.2	Entropy–96.8
3	Random Forests	Gini - 99.0	Gini - 96.4
		Entropy - 99.4	Entropy- 97.0
4	Perceptron	93.447	91.657

Tab.3 Comparison of Accuracy based on no. of features

S. No	Model	10% test data	30% test data	50% test data
1	Naïve Bayes	89	88.3	88.1
2	Decision Tree	Gini - 98.8	Gini – 98.0	Gini- – 97.96
		Entropy– 98.2	Entropy– 98.07	Entropy– 97.8
3	Random Forests	Gini - 99.0	Gini – 97.9	Gini – 98.04
		Entropy- 99.4	Entropy- 97.8	Entropy– 98.1

Tab.4 Comparison of Accuracy based on test split

Definition of Accuracy Measures:

Precision is a measure of accuracy provided a specific class retrieved from predicting. Precision = diagonal element/sum of relevant column

$$\text{Precision} = tp / (tp + fp)$$

where tp and fp are the numbers of true positive and false positive predictions of **p** for the considered class when the actual value is n

Where precision is a measure of exactness or quality, **recall** is a measure of completeness. Recall is nothing but the true positive rate for the class, and support is the number of samples of the true response that lie in that class.

Naïve Bayes				
Class	Precision	Recall	F1-Score	Support
0	0.95	0.92	0.94	447
1	0.48	0.6	0.54	53
Decision Tree using Gini Index				
Class	Precision	Recall	F1-Score	Support
0	0.99	1	0.99	447
1	0.98	0.91	0.94	53
Decision Tree using Entropy				
Class	Precision	Recall	F1-Score	Support
0	0.99	0.99	0.99	447
1	0.92	0.91	0.91	53
Random Forests using Gini Index				
Class	Precision	Recall	F1-Score	Support
0	0.98	0.98	0.98	447
1	0.86	0.83	0.85	53
Random Forests using Entropy				
Class	Precision	Recall	F1-Score	Support
0	0.97	0.98	0.98	447
1	0.84	0.72	0.78	53

Tab.5 Report on classification of various models

V CONCLUSION

In one or two years from now, early adopters may be able to sum up machine learning capabilities with other technologies and interfaces, and perhaps have a two-way dialogue with the smart machine to discuss and authorize approvals and decisions. A framework for intelligent computer-supported decision resolution through negotiation/ mediation is defined through various machine learning models.

Naive Bayes is a continuous classifier that requires to build a classification by hand. It cannot pick the best features for classification using tabular data. There are methods to adapt it to absolute class prediction however they will answer in terms of probabilities like (A 90%, B 5%, C 2.5% D 2.5%) Bayes can perform quite well, and it does not overfit nearly as much so there is no need to prune or process the network.

Decision Trees will work with classification problems and regression problems. They can predict an absolute value like (red, green, up, down) or a continuous value like 2.9, 3.4, etc., Decision Trees only need a table of data, and they will build a classifier directly from that data without needing any up-front design work to take place. By using a decision tree model on a given training dataset the accuracy improves with more and more splits. It is easily possible to overfit the data unless cross-validation (on training data set) is used.

A random forest selects observations/rows and specific features/variables randomly to build multiple decision trees from and then averages the results. After many trees are built, each tree "votes" or chooses the class, and the class receiving the most votes by a simple majority is the "winner" or predicted class. However, the randomness in feature selection cannot be controlled. Accuracy keeps increasing with an increase in the number of trees but becomes constant at certain point. Unlike decision tree, it will not create highly biased model and reduces the variance.

The Random Forests can only work with **tabular data**. On the other hand, Neural Network can work with many different data types like Tabular data, Images, Audio Data and, Text Data, etc.,

The model integrates varied machine learning models and decision-theoretic techniques to provide enhanced conflict resolution and negotiation support in group problem-solving. As state of the art in Decision Support System (DSS) development advances and as DSSs support increasingly more complicated tasks, such machine

learning techniques will become an indispensable part of decision support systems.

VI REFERENCES

[1] OMG, Decision Model, and Notation 1.1 (2016).

[2] F. Hasi´c, L. Devadder, M. Dochez, J. Hanot, J. De Smedt, J. Vanthienen, Challenges in refactoring processes to include decision modeling, in: Business Process Management Workshops, Lecture Notes in Computer Science, Springer, 2017.

[3] J. Vanthienen, F. Caron, J. De Smedt, Business rules, decisions, and processes: five reflections upon living apart together, in: SIGBPS Workshop on Business Processes and Services, 2013, pp. 76–81

[4] T. Biard, A. Le Mauff, M. Bigand, J.-P. Bourey, Separation of decision modeling from business process modeling using new decision model and notation(dmn) for automating operational decision-making, in: Working Conference on Virtual Enterprises, Springer, 2015, pp. 489–496.

[5] R. Ghlala, Z. K. Aouina, L. B. Said, Bpmn decision footprint: Towards decision harmony along bi process, in: International Conference on Information and Software Technologies, Springer, 2016, pp. 269–284.

[6]https://stackoverflow.com/questions/10317885/decision-tree-vs-naive-bayes-classifier

[7]https://www.k2.com/blog/machine-learning-business-decision-making

[8]https://blog.frankfurt-school.de/wp-content/uploads/2018/10/Neural-Networks-vs-Random-Forests.pdf