First-Order Theorem Proving using Classification

Shashwat Chahal

Guru Gobind Singh Indraprastha University

B.Tech AIML

Instructor*: Prof. Amit Choudhary*

Date: 18-05-2023

Basic Information

Title of Project: First-Order Theorem Proving using Classification

Student Name: Shashwat Chahal

Enrollment Number: [redacted]

Email ID: [redacted]

Contact Number: [redacted]

Google Drive Link: [redacted]

Google Website Link: [redacted]

YouTube Video Link: [redacted]

Note: The links provided herewith might be empty or non-responsive yet, they will soon be updated with relevant information.

Title: First-Order Theorem Proving using Classification

1. Introduction:-

The goal of this research project, which focuses on first-order theorem proving, is to create a classification model that can predict which of five potential heuristics would produce the fastest proof. The UCI Machine Learning Repository provided the dataset for this study, which includes details about theorems and the relevant heuristics. The goal is to create a model that can correctly categorise the best heuristic for theorem proving.

2. Methodology:-

The appropriate heuristic for theorem proving will be predicted by the project using classification methods. The actions listed below will be taken:

a. Data Pre-processing: The dataset will go through preparation procedures, which may include managing missing values and feature selection if necessary.

b. Feature engineering: From the dataset, pertinent features will be taken out in order to describe the properties of the theorems and the heuristics that go along with them. This can entail adequately scaling and converting the data.

c. Classification Models: Several classification techniques will be used to train and assess the performance of the model, including logistic regression, decision trees, and random forests.

3. Data Pre-processing:-

In order to handle any missing values and guarantee data quality, the dataset will be pre-processed. To extract useful data from the dataset, including the qualities connected to theorems and their relevant heuristics, feature engineering approaches will be used.

4. Model Development and Evaluation:-

Utilising the proper machine learning libraries, classification algorithms will be put into practise. A portion of the dataset will be used to train the models, while the remaining samples will be used for evaluation. The performance of the models will be assessed and compared using performance indicators like recall, recall interval, recall accuracy, and F1-score.

5. Conclusion:-

The goal of the study is to create a classification model for first-order theorem proving that forecasts the best heuristic for producing the quickest proof. The study will use a variety of classification algorithms to see which approach delivers the maximum accuracy in predicting the right heuristic. This study can help to increase the effectiveness of first-order provers and optimise theorem proving procedures.

References:

1. Dua, D., & Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

2. Quinlan, J. R. (1996). Bagging, boosting, and C4.5. In ACM SIGKDD Explorations Newsletter, 2(2), 25-27.

3. Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.