T22
AI/ML
2019-05-02 15:00

Testing Large Data Sets with Supervised Machine Learning

Presented by:

Alireza Razavi

TD Bank

Brought to you by:

Alireza Razavi

Alireza Razavi has more than 20 years experience as a QA director, Senior Advisor, and Practice Lead in organizations like; TD BANK, BNP PARIBAS BANK, National Bank of Canada, Montreal exchange. He has managed, defined, established norms, standard, best practices and governance strategy for different testing teams both on and off shore. Alireza has experience with test data management and information management, DevOps, continuous integration and AI best practices for testing and has worked on projects using BI, ETL and big data testing management, automation test management, AI automation frameworks, and performance test management.

AI Testing Framework :
Supervised Machine Learning

Alireza Razavi

Agenda

- **Insurance Rating**
- **Supervise Machine Learning**
- **Decision Tree Algorithm**
- **AI Framework**
- **Benefits**
- **Q&A**

Insurance pricing (Rating)

- A *rate* "is the price per unit of insurance for each exposure unit, which is the unit of measurement used in insurance pricing"
- The premium rates set by insurance companies involve calculation methods that incorporate the costs of insuring a person or business
- Insurance pricing methods can vary in terms of the types of variables considered when determining pricing rates. Methods used may consider risk factors, probability factors and individual claims histories depending on the type of insurance involved.
- Every year, price rates should be changed based on different regulation. Hence after each price rate change, the new price rate has to be tested for huge number of coverage data to make sure that insurance premium is correct based on the coverage.
- Testing insurance premium for such a coverage data as input and its variation is almost impossible for any testing team.
- This presentation will present an AI automation testing framework (Supervised machine learning) which is designed and used to solve this issue. A machine Learning algorithm (Decision tree) is used in regression testing

- *Rejda, George; McNamara, Michael, Principles of Risk Management and Insurance (12th ed.). Upper Saddle River, NJ: Pearson Education.*

Supervise Machine Learning

- **Definition:**

  Supervised machine learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs

- **Scope:**

  It infers a function from *labeled* training data consisting of a set of *training examples*. In supervised learning, each example is a *pair* consisting of an input object (typically a vector) and a desired output value.

  A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples.

  An optimal scenario will allow for the algorithm to correctly determine the class labels for unseen instances.

- Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) *Foundations of Machine Learning.* The MIT Press ISBN 9780262018258

Decision Tree Algorithm

- **Decision tree learning** uses a decision tree to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves).

- It is one of the predictive modelling approaches used in data mining and machine learning.

- Tree models where the target variable can take a discrete set of values are called **classification trees**; in these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels.

- Decision trees where the target variable can take continuous values (typically real number) are called **regression trees**

- Quinlan, J. R., (1986). Induction of Decision Trees. Machine Learning 1: 81-106, Kluwer Academic Publishers
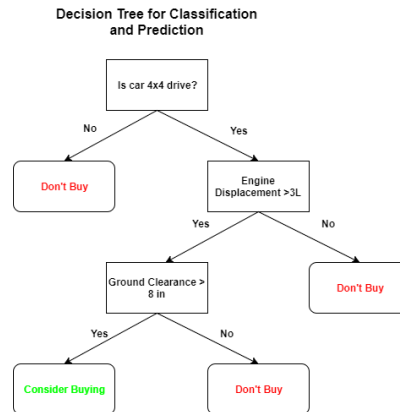
Decision Tree Algorithm 2

- Determine the type of training examples.
- Gather repressive training set
- Determine the input feature representation of the learned function.
- Design corresponding learning algorithm.
- Run the learning algorithm on the gathered training set: determine certain control parameters.
- Evaluate the accuracy of the learned function (separate from the training set)

- Mehryar, Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) *Foundations of Machine Learning*, The MIT Press ISBN 9780262018258

Decision Tree Algorithm 3 (Example)

Decision Tree for Classification and Prediction

Haiban Lan, Decision Trees and Random Forests for Classification and Regression pt.1

AI Framework

- A machine Learning algorithm (Decision tree) has been used in software testing.
  - Detect proper navigation with respect to supplied test data
  - Execute sub-iterations for applicable components to complete desired transaction
- High Level Architecture Diagram
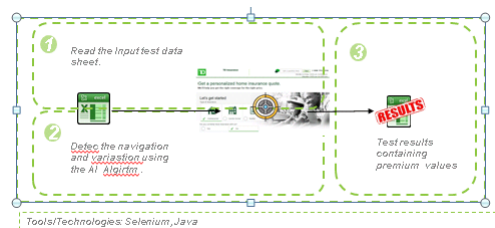
Benefits

- A machine Learning algorithm (Decision tree) is used in software testing for regression testing.
- Test all test cases with all variations
- Improve quality
- Less cost and resource
- Improve time to market
- This framework will optimize data regression testing and regression testing