Jiandong Li

Northeastern University   **Jiandong.li.92@outlook.com**

Personal Detail

Birthday: 1992/8/1   Gender: Male
Native Place: Linyi City, Shandong Province
Hobby: Photography, Table Tennis, Travel, Bicycle Riding

Education Work

March 2017 – now Beijing Huawei Digital Technologies Co., Ltd Level 14
Sept. 2015 – July. 2016 Hosei University of Japan GPA: 3.0
Thesis of Master's Degree of Hosei University A+
Sept. 2014 – Mar. 2017 Northeastern University GPA: 88
Sept.2010 – July. 2014 Yantai University Average GPA: 3.4

Paper

1: Jiandong Li, Atsushi Sato, Runhe Huang and Dongming Chen, "A Rule-based Knowledge Discovery Engine Embedded in Retail Semantic Graph Knowledge Repository", IEEE CBD 2016, August 13th-16th, in Chengdu China.
2: Jiandong Li and Runhe Huang, "Enhanced KStore with the Use of Dictionary and Trie for Retail Business Data", IEEE ICBDA 2016, March 12th-14th, in Hangzhu China.

Awards

Huawei Personal Golden Network Award in 2017
Grant for Education and Research of Hosei University in 2016
Grant for Education and Research of Hosei University in 2015
First Level Scholarship of Northeastern University in 2015
First Level Scholarship of Northeastern University in 2014
Provincial Excellent Graduate of Shandong Province in 2014
School Level Scholarship of Yantai University in 2012
Work Completion Award of Qilu Software Design Competition in 2012

Skill

Data Storage:
- Be familiar with MongoDB CRUD operations and know MongoDB aggregation operations by either the mongo shell or Python.
- Filter data at initial data reading stage by pattern matching strings in queries and by usage of the computed results returned from MongoDB aggregation operations.
- Do precomputation by MongoDB aggregation operations to reduce total data analysis time or to improve time performance.
- Data backup.

Data Preprocessing:
- Filter data by common statistics, such as, data length, max/min value, abs difference, initial value, etc.
- Data resampling or missing data handling by linear interpolation.
- Time alignment for time series.
- Be familiar with using Pandas package to do data wrangling such as clean, transform, merge and reshape by Python.

Data Analytics:
- Time series trend detection, such as, trend growth, surge trend, oscillation trend, (Moving Average Algorithm + 3σ rule or SAX Algorithm).
- Algorithm parameters loading, adjustment, optimization by configuration txt file (json format).
- Time series symbolic representation, (SAX algorithm) .
- Time series regression or forecasting, (ARMA + AR algorithm).
- Time series clustering, (DBSCAN algorithm + k-mediod algorithm).
- Labelled data collecting and self-learning platform construction, (Feature extraction + SVM + RF + cross-validating + feature selection).
- Random-Forest based feature selection.
- Voting strategy for decision system.
- Analysis result backup.
- The interaction between Python and Spring Websites.
- Analysis result visualization, (line chart + bar graph + box diagram (show data distribution))
- Figure Base-64 encoding and decoding.
- Multiple-factor correlation analysis, (PCC + Spearman rank correlation)

Rule Reasoning:
- Be familiar with Jess Rule Engine and can write some simple forward and backward reasoning chaining by Jess.

Master's Research

In today's competitive retail environment, the age that retail companies make decisions based on gut feeling has gone. They must be able to take advantage of all available data to do advanced analytics which can deliver useful insights and guide their decision making to occupy a chunk of market share against competitors. To support data analytics, KID (data-information-knwoledge) model is proposed by prof. Huang and Mr. Sato. KID model is a general from data to knowledge cognitive model which can accumulate experience and gain knowledge by continuously perceiving data, interpreting data into meaningful information, absorbing incoming information, and updating knowledge as humans do. My work is to design and implement the knowledge repository of KID model which should be designed to support interpreting meaningful information, assimilating meaningful information, and updating knowledge. So based on enhanced KStore which is an enhanced version of KStore in data access and google knowledge graph, a two-layers knowledge store prototype has been proposed by us for KID model. KStore is a computer data structure proposed by Jane Campbell Mazzagatti based on the Phaneron of C. S. Peirce and is designed as a storage engine to support business intelligence data storage, queries and analysis. However, after our further research and experiment on this two layers knowledge store model, we found that it had difficulties in updating knowledge and reasoning tasks. The idea that models knowledge repository as a graph to represent concepts and their relationships in retail domain is inherited by us to design and develop a novel knowledge repository for our KID model. Based on Grüninger and Fox method, Neo4j graph database and retail ontology designed by Maryam Fazel Zarandi, a retail semantic graph knowledge repository embedding a pragmatic rule-based knowledge discovery engine is proposed and developed by us for KID model. We also use RFM model to interpret customer data and purchase record data into customer value and assimilate customer value into our graph knowledge repository and deduce answers to retail queries from it with the use of Jess rule engine.