Data mining project proposal

Introduction

Data Mining has also a variety of research fields including Text Mining, Temporal Mining, Stream Mining, Spatial and also in Geographical Mining, Utility Mining, Web mining, Distributed Data Mining, Ubiquitous Data Mining, Hypertext and also Hypermedia Data Mining, Multimedia Data mining, Time Series Data Mining, also in Constraint Based Data Mining, Phenomenal Data Mining etc.

Current Trends in Data Mining

- Software engineering with Data Mining
- Visual Data Mining
- Interactive and scalable methods also for Data Mining
- Application Exploration
- Biological Data Mining
- New methods also for complex Data Mining
- Standardization of query language also in Data Mining
- Multi database and also Multi Rational Data Mining
- Information Security and also Privacy Protection in Data Mining
- Big Analytics integrated also with Cloud Computing

    Hadoop

    MapReduce

    Apache Spark

    Amazon EC2 and S3

Steps in Data Mining

- Understanding of the application/relevant prior knowledge
- Make target set also for discovery
- Data pre-processing and also cleaning
- Reduce invariant representations and also number of data variables
- Select any of the following data mining tasks

    Regression

    Clustering

    Classification

Association Rules

Data visualization

Feature Extraction and Selection

Anomaly Detection

Statistical data analysis

Multidimensional analysis

- Apply data mining algorithm
- Patterns searching
- Knowledge discover

Specific Models Used in Data Mining

- Decision Tree
- Non-negative Matrix Fabrication
- K-means and also O-clustering
- Naïve Bayes Algorithm
- Support Vector Machines
- Apriori and Hashing Techniques
- Neural networks and also expert systems
- Intelligent agents
- Soft Sets also for Machine learning and Data mining
- Genetic Algorithms
- Artificial Neural Networks

Prominent Data Mining Research Topics

- Data integrity, privacy and also security issues in data mining
- Mining of Multi-agent data also using data mining Concepts
- For data mining a unifying theory can also be created
- For network setting data mining can be also used
- High dimensional data and high speed data can be also streamed
- Mining of sequence information and also time unbalanced data
- Distributed data mining applications