List of Topics in Data Mining (last 3 years from major conferences)

Allaparthi Sriteja,
201302139.

Introduction:

In this document, the topics from last 3 years of major datamining conferences ICDM, KDD ... were taken. Based on the category of the topic all these topics were categorised. These clusters were given names according to the list of the topics under it. From this a 2 level hierarchy (topics and subtopics) is formed based on the publications in the recent years in the field of Datamining.

Intermidiate work:

These topics were made into clusters based on the similarities of the work required in the topic. Like this all the sub-topics according to their meanings and similarities in the work are categorised into Clusters accordingly. This will give the 2-level hierarchy of the topics and sub-topics from the recent publications in Data Mining.

The explanation for some of the keywords were given below, Based on this these were divided and arranged under the main topics.

Bag of words :

In this the words were visualised in clusters in the Euclidean space. This model does not give any information regarding the semantics of the sentence. This is the major limitation of this model. So, this comes under Clustering.

Recommender System :

In this, based on the user activity in that domain this system will generate recommendations for individual users.

Decision Tree :

A decision tree is a classification algorithm that automatically derives a hierarchy of partition rules with respect to a target attribute of a large dataset. So, this comes under Classification.

Association Rule :

An Association Rule is a common knowledge model in data mining that describes an implicative cooccurring relationship between two disjoint sets.

Text Mining :

It is a competitive statistical technology to extract relevant information from huge textual unstructured databases. So, the Text analysis, text classification, text analysis comes under this.

Feature Selection :

It is a dimensionality reduction technique that aims to select a subset of the original features of a dataset which offer the most useful information. The benefits of feature selection include improved data visualisation, transparency, reduction in training and improved prediction performance. Feature Analysis, feature Classification, feature extraction come under this.

Support Vector Machine:

It is a machine learning technique that has generated a lot of interest in the pattern recognition community in recent years. It is used to make classification and regression on data from unknown probability distribution.

Data Analysis :

Analysis of data is a process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making. So, data collection, data cleaning, data processing comes under this.

Social Media Mining :

Social Media Mining is the process of representing, analyzing, and extracting actionable patterns from social media data. It deals with the tools to formally represent, measure, model, and mine meaningful patterns from large-scale social media data.

Hierarchy of Topics:

Recommender System

Collaborative Filtering
Content based Filtering
Information Retrieval
E-commerce, Movies
Information Filtering
User preferences

Clustering

Data Clustering
Document Clustering
Fuzzy C mean custering
K means

Association Rule Mining

Market Basket Analysis
Frequent Pattern
Interestingness Measure

Frequent Itemsets
Generalised Association Rules
Fuzzy Association Rules

Support vector Machine

Feature Extraction
Feature Selection
Pattern recognition

Feature Selection

Feature Space
Feature Extraction
Classification Accuracy

Data Stream Mining

Continous Query
Query processing
Sliding window
Stream processing
concept drift

Classification

Decision Tree
Feature Selection
Image Classification
Knn
Text Classification
Support Vector Machine
Feature Extarction

Frequent Pattern Mining

Frequent Patterns
Association Rule Mining
Sequential Patterns
Frequent Itemsets
Pattern Discovery

Text Mining

Document Clustering
Information Extraction
Information Retrieval
Text Classification
Text Clustering
Web mining
Text Analysis

Data Analyis

Data Collection
Data Processing
Microarray Data
Statstical Methods
Time series
Data Cleaning

Social Media Mining

Semantic Sentiment Analysis
Online Social Network
Collective Intelligence
User generated Content
Recommender System