Chapter 1

Introduction

1.1 Overview

For the past two decades, the rate at which World Wide Web has expanded is enormous. Instantaneous access to a large diversity of knowledge items became possible due to the advent of the first generation of the World Wide Web (WWW). The second generation of the WWW is typically denoted by Web 2.0. It has instigated a fundamental change in the way people interact with and through the World Wide Web. This is why Web 2.0 is also referred to as the participatory Web. It can be characterized as a paradigm that facilitates communication, interoperability, user-centered design, information sharing and collaboration on the Web (Sharma Eijlander and Bogers).

Moreover, in the transition to Web 2.0, a paradigm shift from local and solitary to global and collaborative is exhibited. The shift also accords with a shift from accessing and creating information to understanding information and the people who deal with this information. Information management and access has been moving to many distributed places on the Web, instead of, creating, storing, managing, and accessing it on only one specific computer or browser. Collaboratively created websites such as Wikipedia are accessed and edited by anyone from anywhere. Users have started documenting and sharing many aspect of their lives online using blogs, social networking sites, and video and photo sharing sites (Eijlander and Bogers).

A global information society with a growing number of users around the world has emerged due to these developments. This has resulted in to avalanche of informa-
tion at our doorstep. The difficulty in finding what we want, when we need it, and in a manner which best meets our requirements is rapidly and continuously increasing. These days, for users, too many options to choose from is a common and constantly confronted situation. Users need help to explore and filter out their preferences from the myriad possibilities. Internet Search Engines, designed originally to be helpful, now commonly find many thousands of potentially relevant sites, thus losing their effectiveness (Montaner Rigall et al.).

A great amount of research on how Artificial Intelligence (AI) can help people to find out what they want on the Internet has been carried out by the AI community. As a result, users who require assistance in searching, sorting, classifying, filtering and sharing the vast amount of information now available on the Web have widely accepted the idea of recommender systems. The key task of a recommender system is to find items, information sources and people related to the interest and preferences of a single person or a group of people (Montaner Rigall et al.).

1.2 Problem Statement

“Enhancing Performance of Recommender Systems” deal with improving performance of recommender systems applicable to various domains. Goal of this work is to make recommendation methods more accurate and applicable to broader range of real-life needs.

Different applications generate different categories of data. Depending on nature of data, data processing varies. Suitable representation of data would also optimize performance of a recommender system.

Hence, need is to identify suitable data processing scheme along with the best suited data representation to optimize performance of recommender systems pertaining to different domains.

1.3 Organization of the Thesis

Rest of the thesis is organized as follows:

Chapter 2 (Background): This chapter presents an overview of the field of recommender systems. Three main recommendation approaches along with current generation recommender systems are discussed. However, literature survey specifically
related to different recommender systems developed in this thesis, is discussed, whenever the corresponding chapter is elaborated.

Chapter 3 (Predicting Direction of Movement of Stock Price and Stock Market Index): This chapter deals with predicting movement of stock price and stock market index. Four prediction models, Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest (RF) and naive Bayes (NB) with two approaches for input to these models are compared. Experiments are carried out on 10 years of historical data of two stocks and two stock market indices from Indian stock markets.

Chapter 4 (Predicting Stock Market Index using Fusion of Machine Learning Techniques): This chapter addresses the task of predicting future values of stock market index. A two stage fusion approach is proposed and three hybrid models are developed for the task. The chapter also discusses about experimental results and compares proposed fusion models with conventional single stage models.

Chapter 5 (Social Resource Recommendation using Learning from Positive and Unlabeled Examples): The problem of recommending social resources is focused in this chapter. Social resources are part of social resource sharing websites. Resources on these websites fall in to the category of positive and unlabeled data. Two step techniques and a direct method capable of learning from positive and unlabeled data is proposed in the chapter for the task of social resource recommendations.

Chapter 6 (Recommending Tags for new Resources in Social Bookmarking System): A tag recommendation algorithm based on the concept of multi-label classification is presented in the chapter. Impact of feature selection and representation of the resource on the performance of the recommender is also presented.

Chapter 7 (Movie Recommender System - Hybrid Filtering Approach): The focus of this chapter is movie recommendation task. Prediction task is modelled as classification task where the aim is to predict whether the movie will be liked or disliked by the user. An item based recommender which combines usage, tag and movie specific data such as genres, star cast and directors is proposed in this chapter.

Chapter 8 (Evaluating a Recommender learnt from Labeled and Unlabeled Data): The main objective of this chapter is to examine the influence on the accuracy of the recommender when it is built using unlabeled examples in addition to the labeled examples. Co-Training algorithm which allows to incorporate unlabeled
examples while learning a classifier/recommender is discussed. Usefulness of this algorithm is investigated by means of experimental study using hetrec2011-movielens-2k data set.

Chapter 9 (Conclusions and Future Work): In this chapter, conclusions drawn from various implementations mentioned above are discussed. Future scope of work in this area is also mentioned in the chapter.

A separate section for the Indexes used in the thesis is covered towards the end.

The Works Cited section consists of related research work cited in the thesis work.