1 METHODOLOGY

This chapter

1.1 The Problem

Like any other program this bike sharing service also suffers few major problems. Bike rebalancing problem being the most important and significant of them all. This rebalancing problem occurs when a bike station is either full or empty and makes it difficult for users to rent or to park their rented bike. Thanks to the current prediction systems, it is possible to predict the usage of the bike share program and therefore making the users to have better experience of the bike share system. The objective is to make sure the users can walk to a bike docking station and find bikes to rent and also to discharge the rented bikes without having to wait for another user or going to another station to do that.

In order to find the problem and to set the goal of my thesis, first I did some statistical analysis on the bike share data and from those studies I saw that the distribution of the data is not balanced and same goes for gender of the customers. Therefore, I tried to find a solution to overcome this two problem which is to forecast hourly rental bike demand for both men and women. I treated this problem as a regression forecasting with time series features and proposed Convolutional Neural Network (CNN) to do the forecasting which outperforms other traditional machine learning methods. I have used the data obtained from the City Bike system and weather data to forecast hourly rental demand of the Bike share system.

1.2 Datasets

The goal of this thesis is to forecast bike demand using the users historical bike usage patterns along with the weather data and userâĂŹs information in a time series arrangement. Therefore, this thesis attempts to answer the following questions âĂŞ How many bikes are needed every hour? Who are renting those bikes? How weather is affecting the bike share usage? Is it a holiday, weekend or normal weekday?. Therefore, datasets should contain three major information: 1) bike renting time and details, 2) users information, 3) weather data as well as 4) Holiday/Weekday-Weekend information.

Collected dataset on the city bike share system contains details about the trip and the user who is renting it from JanuaryâĂŹ2015 to JulyâĂŹ2015. I have also collected

weather datasets during that period and used JanuaryâĂŹ2015 to JuneâĂŹ2015 data sets to create training set.

figure ??datasets shows the overview of the dataset preparation to build the forecasting model. Detailed information about the dataset extracting and processing will be discussed later in this chapter.

figure adjustboxaddcode=minipage Dataset Overview. datasets