Department of Information Engineering and Computer Science

Bachelor's Degree in Informatica

Final Dissertation

Title

Subtitle (optional)

Supervisor
......

Student
......

Academic year 2019/2020

Contents

Summary

Sommario è un breve riassunto del lavoro svolto dove si descrive l'obiettivo, l'oggetto della tesi, le metodologie e le tecniche usate, i dati elaborati e la spiegazione delle conclusioni alle quali siete arrivati.

Il sommario dell'elaborato consiste al massimo di 3 pagine e deve contenere le seguenti informazioni:

- contesto e motivazioni
- breve riassunto del problema affrontato
- tecniche utilizzate e/o sviluppate
- risultati raggiunti, sottolineando il contributo personale del laureando/a

1 State of art

This chapter gives an overview of the available methods of time series forecasting and queue analysis.

1.1 Terminology

- *Inflow/outflow rate*: The number of customers that enter/exit the supermarket in a given interval.
- *Arrival rate*: The number of customers that arrive at the checkouts in a given interval.
- *Service rate*: The maximum number of customers that can be served by each terminal in a given interval.
- *Dwell time*: Time spent by a customer in the supermarket.

1.2 Time series forecasting

A *time series* is a collection of observations made sequentially through time. *Time series forecasting* is the use of a model to predict future values based on previous observations.

A time series can be decomposed into four components:

- *Trend*: Linear/nonlinear increasing or decreasing behavior of the series over time.
- *Seasonality*: Repeating patterns or cycles behavior over time.
- *Noise*: Variability of the observations not explainable by the model.

The following sections describes different techniques that can be used to obtain a forecast of the future values.

1.2.1 Exponential smoothing

In the *Simple exponential smoothing* technique, the forecasted values are based on a weighted average of the previous values, where the most recent observations are given more importance using larger weights [1].

A more complex implementation is the *Holt-Winters exponential smoothing* approach, which is able to decompose the time series into level, trend and seasonal component, giving a more precise forecast. However, this methods are unable to model more complex series with multiple seasonality patterns, as in a supermarket inflow rate.

Taylor's *double seasonal exponential smoothing* method [2] was developed to forecast time series with two seasonal cycles: a short one that repeats itself many times within a longer one.

1.2.2 Artificial Neural Networks

1.2.3 ARIMA models

ARIMA (AutoRegressive Integrated Moving Average) models provide another approach to time series forecasting. Exponential smoothing and ARIMA models are the two most widely used approaches to time series forecasting, and provide complementary approaches to the problem. While exponential smoothing models are based on a description of the trend and seasonality in the data, ARIMA models aim to describe the autocorrelations in the data.[1]

AutoRegressive means the model use the relationship between an observations and some number of lagged previous observations to generate a linear regression model. *Integrated* refers to the use of differencing of raw observations (e.g. subtracting an observation from an observation at the previous time step) in order to make the time series stationary. A stationary time series has mean, variance and autocorrelation constant over time. *Moving Average* means that the model uses the dependency between an observation and the residual error from a moving average model applied to lagged observations.

The parameters of the ARIMA model are defined as follows:

- p: The number of lagged observations included in the model, also called the lag order.
- d: The number of times that the raw observations are differenced, also called the degree of differencing.
- q: The size of the moving average window, also called the order of moving average.

A simple version of the ARIMA model can be used to forecast only non-seasonal time series, but it can be extended to model seasonal patterns (SARIMA, Seasonal ARIMA). However, only a single seasonal effect can be modelled with SARIMA. A solution is to use the SARIMAX model, where the seasonal effect are used with Fourier terms as exogenous variables. This gives a better approximation and other exogenous variables could be considered (i.e. weather) to further improve it.

1.3 Queueing theory

Queueing theory is the mathematical study of waiting lines, or queues. A queueing model is constructed so that queue lengths and waiting time can be predicted.

There are various types of queueing models, but the best representation of a supermarket queue is probably given by the *M/M/c* model, that represent a system with an infinite queue capacity, where the inflow rate follow a Poisson distribution and the service rate of the c servers follow an exponential distribution. The customers are served in FCFS order (First Come First Served).

This model could be used to predict the queue length given the arrival rate and service time distributions. The main issue with this analytical approach in the context of supermarkets is that the arrival time and service time cannot be expressed with a probability distribution with constant mean, since this rates change greatly between different hours and days. The formulae provided can still be used if the measurements are in discrete intervals.

1.4 Literature review

There are various studies that try to analyze and predict the queue length in different settings.

1.5

2 Data analysis

This chapter describes the analysis

2.1 Time series analysis

3 Methodology

This chapter describes the methods used to build the prediction model.

3.1 Model overview

The final model is composed by different parts that cooperates to obtain the final result.

First, the measured inflow rate for the recent past is combined with a forecast of the inflow rate for the immediate future. This values, combined with a dwell time prediction, are used to obtain a prediction of the arrival rate at the checkouts. The arrival rate is then used with queueing theory methods to calculate the expected queue length for a given interval.

3.2 Inflow rate forecast

The inflow rate can be expressed as a time series, hence time series forecasting methods were used. As described in Chapter 1.2, different forecasting methods were tested and the best was selected based on the performance obtained.

3.2.1 Persistence model

The first model tested was used to define a baseline in performance. This gives an idea of how well the other models could perform on the time series, and if a model's performance are worse than this baseline it should no be considered. The baseline was obtained by a *persistence model*, were the last measured value y_{t-1} is used as forecast:

$$\hat{y}_t = y_{t-1}$$

3.2.2 Naïve model

As seen in Chapter 2, the arrival rate time series presents a strong seasonal component, so the forecast should be based on this repeating pattern. The forecast value was calculated as the average of the values in the previous weeks for the same day and time:

$$\hat{y}_t = \frac{1}{N} \sum_{i=1}^{N} y_{t-iW}$$

where W is the number of intervals in a week and N is the number of previous week considered.

The model performance were worse than the baseline. Looking at the results, it is clear that this model is not able to consider the variations in the inflow rate that can be caused by external factors, e.g. holidays. To take in account this variations, a local drift was calculated using the forecasted errors value of the immediate past:

$$\hat{y}_t = \frac{1}{N} \sum_{i=1}^{N} y_{t-iW} + \frac{1}{M} \sum_{j=1}^{M} \frac{1}{N} \sum_{i=1}^{N} y_{t-iW-j}$$

3.3 Arrival rate forecast

3.4 Queue length forecast

4 Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam

4.1 Pellentesque

5 Conclusions

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec sed nunc orci. Aliquam nec nisl vitae sapien pulvinar dictum quis non urna. Suspendisse at dui a erat aliquam vestibulum. Quisque ultrices pellentesque pellentesque. Pellentesque egestas quam sed blandit tempus. Sed congue nec risus posuere euismod. Maecenas ut lacus id mauris sagittis egestas a eu dui. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Pellentesque at ultrices tellus. Ut eu purus eget sem iaculis ultricies sed non lorem. Curabitur gravida dui eget ex vestibulum venenatis. Phasellus gravida tellus velit, non eleifend justo lobortis eget.

5.1 Pellentesque

Bibliography

[1] Athanasopoulos G. Hyndman R. J. *Forecasting: Principles and Practice*. OTexts, 2nd edition, 2018.

[2] James W. Taylor. Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19(4):715–725, 2003.