Forecasting Bike Sharing Demand Using Fuzzy Inference Mechanism

Syed Moshfeq Salaken(✉), Mohammad Anwar Hosen, Abbas Khosravi, and Saeid Nahavandi

Center for Intelligent Systems Research, Deakin University, Geelong, Australia
syed.salaken@gmail.com

Abstract. Forecasting bike sharing demand is of paramount importance for management of fleet in city level. Rapidly changing demand in this service is due to a number of factors including workday, weekend, holiday and weather condition. These nonlinear dependencies make the prediction a difficult task. This work shows that type-1 and type-2 fuzzy inference-based prediction mechanisms can capture this highly variable trend with good accuracy. Wang-Mendel rule generation method is utilized to generate rulebase and then only current information like date related information and weather condition is used to forecast bike share demand at any given point in future. Simulation results reveal that fuzzy inference predictors can potentially outperform traditional feedforward neural network in terms of prediction accuracy.

1 Introduction

Bike sharing demand forecasting in city level granularity is a relatively new problem in research community, with only a few researchers approaching this problem with machine learning techniques [1,2]. In addition, most of the existing works in current literature addresses the problem based on short span of prior data for knowledge discovery. Only [1–3] utilizes approximately 2 years of data for prediction. However, authors in [2] used historical data for bike share demand forecasting ahead of 24 hours, researchers in [1] used this for anomaly detection and authors in [3] used this for forecasting with signal processing techniques. Our work utilizes the same dataset used by [1,2]. Among other related works, researchers in [4] addresses balancing uneven demand among stations with several mover trucks, authors in [5] addresses the same issue with only one mover truck, authors in [6] investigates availability of bike in station within a predefined time span in the city of Dublin and researchers in [7] addresses the near future bike availability prediction in stations. Work in [8] forecasts the number of bikes entering and exiting a particular station with neural network as a part of building fuzzy and neural network coupled relocation algorithm for bike sharing system. Please note, [8] does not use fuzzy logic for bike demand forecasting. Finally, researchers in [9] addresses the bike availability issue with a small 7 week dataset.

S. Arik et al. (Eds.): ICONIP 2015, Part III, LNCS 9491, pp. 567–574, 2015.
DOI: 10.1007/978-3-319-26555-1_64

Among these works, only investigation in [3] forecasts the bike share demand with current variables like weather, holiday and number of users. However, even this work does not attempt to forecast the bike share demand based on only date, holiday and weather without the knowledge of user base volume. Finally, none of the works in existing literature employs a solution which not only makes a decent forecast, but also has the capability to explain the reason behind the forecasted volume.

In this work, we demonstrate the fuzzy inference can predict the bike share demand with good accuracy at any given point of time with the knowledge of calendar information coupled with weather variables like temperature, wind speed and humidity. Once the fuzzy inference model is built with prior knowledge, past data are no longer required to make a prediction. This is the key difference of our work with [7], which is most similar to our work based on a similar dataset from Barcelona. However, work in [7] forecasts only station level demand, not system level demand and therefore, our results are not comparable with theirs. Research in [2] predicts one day ahead demand by introducing several artificial time dependent features and missing samples. This kind of data preprocessing is not necessary for fuzzy inference. In addition, this fuzzy inference mechanism has the ability to explain the causal effect of date and weather situation on the prediction value.

Please note, no new fuzzy inference mechanism is claimed in this work, instead it is demonstrated that existing fuzzy inference mechanism, including type-1 and interval type-2 fuzzy sets, is capable of capturing the highly diverse nature of bike share demand. To the best of our knowledge, this is the first fuzzy inference based prediction study on bike share demand to forecast overall bike sharing demand. It is also shown that fuzzy inference mechanism is better than neural network prediction on the same dataset, without performing any optimization on either FLS parameters or neural network parameters, and therefore provides a better overall accuracy along with easily interpretable prediction mechanism. This, of course, does not imply anyway that a better tuned neural network will not outperform a FLS. The focus of this work lies within the scope of demonstrating that a simple FLS can forecast the highly variable nature of bike share demand while offering an insight to the causal relationship among demand and different contextual factors.

The rest of the paper is organized as follows: Sect. 2 describes the experimental procedure along with preprocessing, Sect. 3 discusses and compares the results, and finally Sect. 4 concludes the work.

2 Experimental Procedure

The data used in this study is taken from Capital Bikeshare system, Washington D.C., USA for year 2011 and 2012, as provided in [1]. It is available in daily and hourly format. Both of them contain index, date, season, year, month, day of the week, hour (only in hourly format) weekend, public holiday, weather situation, temperature, feeling temperature e.g. real feel of the temperature, humidity and wind speed. It also provides number of registered and casual users in addition to total user count.

2.1 Data Preparation

Original data contains few extraordinary environmental scenario which is not helpful to this regression problem, but can be of significant importance for some cases. As these are not in the scope of this work, those data are discarded as part of data cleansing. For example, effects of hurricane Sandy and extra ordinary rain on December 7, 2011 are discarded. After this is done, date information and index are also discarded. This is done to ensure the purpose of this work, i.e. showing fuzzy inference is capable of forecasting highly diverse nature of bike sharing demand, is not biased by gradually increasing demand of the service. Next, the data is shuffled randomly before split into training and testing chunks. This randomization may provide different level of success in forecasting. Therefore, each forecasting computation is run 5 times to make the result statistically meaningful. Data is neither scaled nor normalized in this study. However, other aspects of developing fuzzy inference model for this purpose will be undertaken in future work. Please note that, one out of every six samples is taken into consideration for hourly forecast to reduce the forecasting time. If all data is taken into account, it is very likely to produce even more accurate result.

2.2 Fuzzy Rulebase Construction

This work utilizes the renowned Wang-Mendel rule generation method [10] to generate fuzzy rulebase from numerical data. As different input variables are known from historical data, it is very easy to produce an exhaustive rulebase from them. After which this database is pruned by picking only one rule from a conflict group. Gaussian membership function is utilized throughout the work to ensure whole input domain is covered in order to avoid any abrupt change in control surface. A total of 10 membership functions are used on each of 9 inputs to determine the fuzziness of each input sample. Because the range of output is very large, 35 membership functions with a very large standard deviation are used to capture the output domain.

Two types of fuzzy inference mechanism are used and compared. Please note, the parameters of fuzzy inference system or fuzzy logic system are not optimized as this work only intends to show fuzzy inference mechanism is a good predictor for this kind of data. In future works, tuning will be done to achieve best performance form fuzzy inference process. Parameter values used in this experiment are listed in Table 1. Sample membership functions for both type-1 (T1) and interval type-2 fuzzy inference (IT2 FLS) are shown in Fig. 1a and b. MFs on all input are not shown due to page limitation.

2.3 Fuzzy Inference

Once the fuzzy rulebase is constructed, this is applied on testing data. This is done by taking one sample of data which contains one instance of every input and passing that to fuzzy inference function. Inside the inference mechanism, fuzziness of each input is first done by the means of membership degree computation. Afterwards, the applicability of each rule is calculated and center of corresponding rule output is captured. Finally, they are aggregated and defuzzified using the method described in [10]. For interval type-2 fuzzy inference, first the firing interval is calculated in the same way applicability of rule is computed in type-1 inference and then defuzzification is done using EIASC algorithm [11].

2.4 Rulebase Adaptation

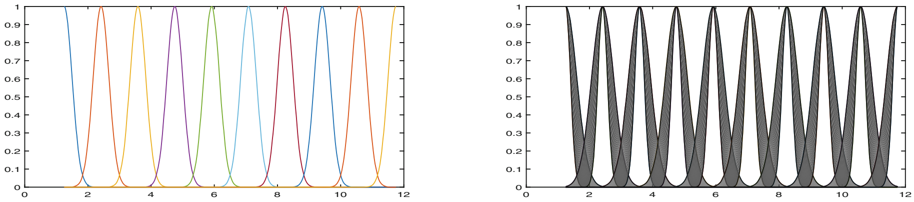
As described in [10], rulebase can be adapted to enhance the performance of fuzzy inference based prediction. A similar approach is taken in this work as well to address the highly variable nature of bike share demand. When one sample of test set is predicted, it is incorporated in fuzzy rulebase with actual output and used to produce another set of exhaustive and pruned rulebase. This effectively increases the exhaustive rulebase size which is used throughout the prediction. In addition, exhaustive rulebase is used in first inference, instead of pruned rulebase, to ensure highest prediction accuracy from the beginning.

2.5 Neural Network Prediction

Since there is no work in literature which tackles this forecasting problem from the perspective we are investigating, a neural network is built, trained and tested for performance evaluation of fuzzy inference mechanism. All inputs and outputs are scaled using their minimum and maximum value prior to neural network training and then final result in scaled back for comparison with actual output. However, no such preprocessing is done for fuzzy inference based prediction. In the neural network, a 3 layer configuration is used with 1 hidden layer and 10 hidden neuron. Bayesian regularization [12,13] is used as training algorithm as it gives more generalization power to neural network [12,14]. Sum of squared error is used as cost function as this is proven to produce most accurate prediction after few trial and error.

Table 1. Parameter values in experiment

Inference type		Input	Output
	Number	9	1
	Range	[0,25]	[22,8714]
T1	Number of MF	10	35
	Standard deviation	0.25	60
IT2	Number of MF	10	35
	Standard deviation	0.15, 0.35	60



(a) MFs on input-03

(b) MFs on input-03

Fig. 1. Samples of T1 and IT2 membership functions on input

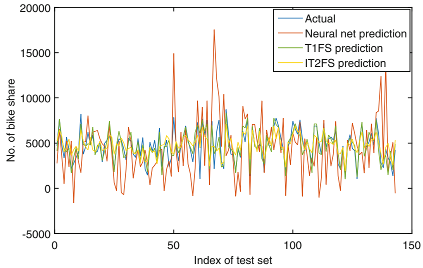
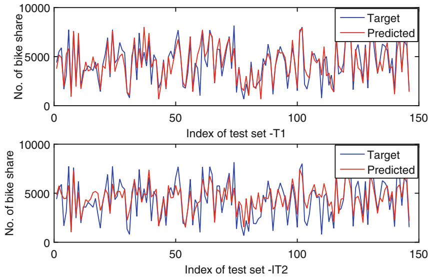


Fig. 2. Comparison of fuzzy inference and neural network prediction (Bayesian regularization, 10 neuron in 1 hidden layer and SSE)

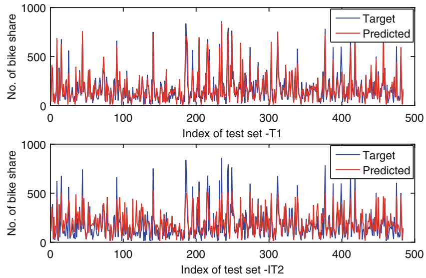
3 Results and Discussion

As RMSE is prone to occasional large error and this dataset contains a wide range of output data (see Table 1), which is neither scaled nor normalized for fuzzy inference, root mean squared logarithmic error (RMSLE) is used as the primary metric for comparison. Because RMSLE is not easy to interpret, other metrics e.g. mean absolute percentage error and percentage of predictions under defined threshold value are also presented. Even though none of these metrics directly shows the level of success in capturing the trend, they establish a benchmark for further study. These values are not comparable with existing literature as this particular problem was never investigated within research community from the same perspective i.e. predicting future demand based on current knowledge only where no previous demand (e.g. demand of last 24 hours etc.) is known. Few related works utilized relative absolute error [1,7], root relative squared error [1], root mean square error for 5 min [6] and mean absolute error on a different, but slightly related, problem [3,9].

As the data provides two sets of datafile for hourly and daily usage, fuzzy inference mechanism is applied on both of them. Result for daily forecasting performance can be seen in Table 2 and Fig. 3a. Hourly forecasting performance is listed in Table 3 and Fig. 3b. From Table 2, note that RMSLE for type-1 (T1) prediction is always lower than interval type-2 (IT2) prediction. This surely indicates that T1 prediction is more accurate than IT2 inference based prediction.



(a) Daily forecasting performance (run-5)



(b) Hourly forecasting performance (run-5)

Fig. 3. Forecast performance

This can be confirmed by respective mean absolute percentage error (MAPE) value as well. However, none of them captures the success of fuzzy inference in capturing the highly variable trend of bike share demand for either T1 prediction or IT2 prediction. Therefore, an acceptance threshold variable is also introduced and number of prediction under this error threshold is calculated in percentage. Note that, this metric shows the trend capturing success partially and preserve the earlier observation of higher T1 prediction success. However, it is neither related with RMSLE nor MAPE value. This can be realized by examining the first and second row of Table 2 which shows a decrease of both RMSLE and MAPE with a decrease of "close point(%)" value whereas an increase is intuitively expected as those error values decreased.

From Fig. 3 and Table 3, observe that fuzzy inference based prediction is even more successful in capturing the changing nature of bike share demand in the hourly forecast. Even in this case, T1 inference based prediction is more accurate than IT2 based prediction. However, please note that no optimization is done on either T1 FLS or IT2 FLS in order to tune their parameter and enhance the performance. It is indeed possible that a better tuned IT2 fuzzy logic system (FLS) may outperform a T1 FLS.

To provide a benchmark, daily forecast prediction of a neural network containing one hidden layer with 10 neurons, which is trained with Bayesian regularization algorithm, is compared with that of fuzzy inference system. The comparison is shown in Fig. 2. Please note, this configuration of neural network is finally chosen after few trial and error with different number of hidden layer neurons and training algorithm as this configuration is found to be producing most accurate prediction. Observe from this figure that neural network prediction is not closely following the target data and occasionally showing very large deviation. Clearly both T1 and IT2 FLS prediction outperforms neural network prediction. Please note, other training algorithms and different number of neurons in hidden layer are also tested and the best result is presented in this figure. However, this result does not imply that a better tuned neural network will not outperform the fuzzy inference based system. Making such a comparison with optimization techniques is beyond the scope of this work and will be investigated in future works.

Table 2. Performance in daily forecast

Run ID	Acceptance threshold	Type-1			Interval type-2		
		RMSLE	MAPE	Close point (%)	RMSLE	MAPE	Close point (%)
1	500	0.3194	24.3735	53.4247	0.3952	36.2882	36.3014
2	500	0.3024	22.3361	48.6301	0.3549	30.626	31.5068
3	500	0.3078	24.5076	47.2603	0.4078	37.2364	34.9315
4	500	0.3694	28.3052	52.0548	0.4431	41.7659	36.3014
5	500	0.3275	24.2118	50	0.3895	35.3197	36.9863

Table 3. Performance in hourly forecast

Run ID	Acceptance threshold	Type-1			Interval type-2		
		RMSLE	MAPE	Close point (%)	RMSLE	MAPE	Close point (%)
1	50	0.4577	36.86	75.4132	0.6596	82.131	57.2314
2	50	0.4922	39.455	72.314	0.6226	75.2026	56.6116
3	50	0.4682	39.2181	71.4876	0.6202	74.5585	53.719
4	50	0.487	39.4747	72.5207	0.6309	77.0695	55.5785
5	50	0.461	40.9236	72.314	0.6351	78.5359	53.0992

Furthermore, it is important to note that results of this study show that T1 fuzzy system outperforms IT2 fuzzy systems in terms of prediction accuracy. This observation complies with existing literature as current literature states that only optimally tuned IT2FLS may outperform another optimally tuned T1FLS [15]. We want to emphasize again that no such optimization is done in this study and footprints of uncertainty for IT2 FS are arbitrarily designed.

4 Conclusion

In this study, it is shown that fuzzy inference mechanism can effectively capture the rapidly changing nature of bike share demand in city level forecasting. Both type-1 and interval type-2 fuzzy inference models are investigated. The result demonstrates that unoptimized fuzzy inference systems can widely outperform unoptimized traditional feed forward neural networks. Thus, simple fuzzy inference based prediction is more effective than costly neural network based prediction approach in the sense of achieving insight among bike sharing demand and contextual factors. Most importantly, this work shows that forecasting of bike sharing demand can be done with fuzzy inference, which has never been done before using a fuzzy logic system in existing literature.

References

1. Fanaee-T, H., Gama, J.: Event labeling combining ensemble detectors and background knowledge. Prog. Artif. Intell. **2**, 1–15 (2013). doi:10.1007/s13748-013-0040-3
2. Giot, R., Cherrier, R.: Predicting bikeshare system usage up to one day ahead. In: 2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS), pp. 22–29. IEEE (2014)

3. Borgnat, P., Abry, P., Flandrin, P., Robardet, C., Rouquier, J.-B., Fleury, E.: Shared bicycles in a city: a signal processing and data analysis perspective. Adv. Complex Syst. **14**(3), 415–438 (2011)
4. Benchimol, M., Benchimol, P., Chappert, B., De La Taille, A., Laroche, F., Meunier, F., Robinet, L.: Balancing the stations of a self service bike hire system. RAIRO-Oper. Res. **45**(01), 37–61 (2011)
5. Chemla, D., Meunier, F., Calvo, R.W.: Bike sharing systems: solving the static rebalancing problem. Discrete Optim. **10**(2), 120–146 (2013)
6. Yoon, J.W., Pinelli, F., Calabrese, F.: Cityride: a predictive bike sharing journey advisor. In: 2012 IEEE 13th International Conference on Mobile Data Management (MDM), pp. 306–311. IEEE (2012)
7. Froehlich, J., Neumann, J., Oliver, N.: Sensing and predicting the pulse of the city through shared bicycling. In: IJCAI, vol. 9, pp. 1420–1426 (2009)
8. Caggiani, L., Ottomanelli, M.: A modular soft computing based method for vehicles repositioning in bike-sharing systems. In: 2012 Proceedings of EWGT2012 - 15th Meeting of the EURO Working Group on Transportation, Procedia - Social and Behavioral Sciences, vol. 54, pp. 675–684, Paris, September 2012. http://www.sciencedirect.com/science/article/pii/S1877042812042474
9. Kaltenbrunner, A., Meza, R., Grivolla, J., Codina, J., Banchs, R.: Urban cycles and mobility patterns: exploring and predicting trends in a bicycle-based public transport system. Pervasive Mobile Comput. **6**(4), 455–466 (2010)
10. Wang, L.-X., Mendel, J.M.: Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man Cybern. **22**(6), 1414–1427 (1992)
11. Wu, D., Nie, M.: Comparison and practical implementation of type-reduction algorithms for type-2 fuzzy sets and systems. In: 2011 IEEE International Conference on Fuzzy Systems (FUZZ), pp. 2131–2138. IEEE (2011)
12. MacKay, D.J.: A practical Bayesian framework for backpropagation networks. Neural Comput. **4**(3), 448–472 (1992)
13. Girosi, F., Jones, M., Poggio, T.: Regularization theory and neural networks architectures. Neural Comput. **7**(2), 219–269 (1995)
14. Foresee, F.D., Hagan, M.T.: Gauss-Newton approximation to Bayesian learning. In: Proceedings of the 1997 International Joint Conference on Neural Networks, vol. 3, pp. 1930–1935. IEEE, Piscataway (1997)
15. Wu, D., Mendel, J.: Designing practical interval type-2 fuzzy logic systems made simple. In: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 800–807, July 2014