

COMPARISONS OF VARIOUS MEASUREMENTS IN WI-FI SIGNAL STRENGTH BY TUNING OF PARAMETERS USING SVM

Dr.Ayyappan.G, Dr.A.Kumaravel

Associate Professor, Machine Learning Group, Department of Information Technology, BIHER, Bharath Institute of Higher Education and Research, Chennai.

Professor, Machine Learning Group, Department of Information Technology, BIHER, Bharath Institute of Higher Education and Research, Chennai.

ayyappangmca@gmail.com

Abstract:

Detecting users in an indoor environment based on Wi-Fi signal strength has a wide domain of applications. This can be used for objectives like locating users in smart home systems, locating criminals in bounded regions, obtaining the count of users on an access point etc. The paper develops an optimized model that could be deployed in monitoring and tracking devices used for locating users based on the Wi-Fi signal strength they receive in their personal devices. Here, we procure data of signal strengths from various routers, map them to the user's location and consider this mapping as a classification problem. In this research work focuses on optimization strategy those results in better accuracy of the model by using the support vector machine.

Keywords: SVM, User localization, Wi-Fi signal strength, Fuzzy logic, Polykernel

I INTRODUCTION

To predict the user's location accurately, a definite and consistent model has to be trained and deployed in a tracking or monitoring device. We measure the Wi-Fi signal strength received from various routers in a bounded location and train the neural network so that it could further predict the user's location for an unknown tuple set having signal strengths. Here, we consider a setup at an office location in Pittsburgh, USA. The office has seven Wi-Fi routers and its signal strengths received from these routers categorize the location of user in the conference room, kitchen or the indoor sports room. Sample data tabulated is shown in Table 1. WS1 corresponds to the signal strength received from the router 1, WS2 corresponds to the signal strength received from the router 2, and similarly for the other routers. The class labels corresponding to the conference room, kitchen and the indoor sport are labelled 1, 2 and 3 respectively. In our setup facility, we have considered an Android device and tabulated strengths of wireless signals captured by the device. At certain locations, the signal strengths were observed by polling the wireless signal strength at a constant time interval (every 1 s considered here). This was again repeated for other locations and suitable data was collected for one thousand and five hundred observations made at this facility for seven different routers. The model developed here, could hence be reused according to the scenario of the bounded location and the number of wireless routers in the physical facility. This data is being formulated into a pattern classification dataset by considering the seven wireless routers as the input dimensions which are used to predict the user's location in an office as one of the three dimensional categories.

In this paper presents section 2 of this paper explains the detail on the related works. In section 3 presents the materials and methods adopted and section 4 presents the details of the experiments and discussions. Finally section 5 concludes the paper by sharing our inferences of this research work.

II RELATED WORKS

In this section focuses on related works of this research work. Advancements in location based services have enabled wide prospects in mobile computing. Many strategies have been adopted to provide users with custom locality based services. These strategies have shown a tremendous boom in e-commerce revenues, embedded smart systems, location based recommender systems and various other fields. Technologies like the GPS, Bluetooth and Wi-Fi could be exploited to provide such services.

Jayant G. Rohra et.al[1]propose the FPSOGSA that overcomes the possibilities of trapping itself in the local minima and enhances the probability of a higher convergence rate. At higher iterations, it gradually decreases

the error rate rather than attaining saturation, as seen in PSOGSA. It obtains better convergence, enhances the ability of optimizing the neural network and hence reducing the mean square error of the Fuzzy Neural Network (FNN). Many such approaches have been used to train User Localization in an Indoor Environment Using Fuzzy Hybrid 287 various classifiers. Artificial Neural Network (ANN) is a model easy to understand and use. More importantly it is nonlinear and non-parametric in nature. ANN is largely used to solve various classification and forecasting problems with the Back Propagation (BP) algorithm. However, the BP convergence is slow and not guaranteed. Therefore, they made optimization strategies to attain faster convergence and higher accuracy rates. Hence, we introduce the hybrid PSOGSA strategy as an optimization strategy here. On the other hand, the ANN is said to be a black box learning approach. It cannot deal with uncertainties. To overcome this, we introduce the fuzzy component. Fuzzy is quite good in handling uncertainties and can also interpret the relationship between the input and output by producing rules. Hence, we introduce the FPSOGSA algorithm.

Salazar et al. [2] introduced methods to predict the behavior of people by monitoring their daily movements. Such location detection systems could also be used in panic situations and disasters, when people require necessary rehabilitation. Nguyen et al. [3] suggests recognition techniques for patients suffering from severe brain injuries who could be monitored by observing patterns in their movements.

Pei et al. [4] proposed SVM techniques and showed better classification rates compared to other existing learning techniques. Cho [5] proposed learning methodologies to categorize the locality of indoor and outdoor regions using the location service logs of smart phones. Zou et al. [6] introduced an indoor localization mechanism based on extreme machine learning strategies and depicted its easy adaptation to versatile environments.

Zadeh [7] introduced the fuzzy set theory that has been widely adopted in many domains like real-time controllers, diagnostic systems etc. Real world data has various dimensions, much more than the classical logic of true or false. The fuzzy logic is used to correlate real life scenarios representing probabilities measuring the degree of truth in the range 0 to 1.

Jang and Sun [8] proposed the interesting concept of modeling neural networks with fuzzy logic and parametrizing control. The neuro-fuzzy strategy alone would not be sufficient to attain the best throughput to the neural networks. The need for our problem lies to detect users at right locations using better learning techniques. But most of the techniques proposed lack the apt usage of optimization strategies that train the model rightly. We look into meta heuristic techniques that promise sufficiently good solutions to optimization strategies.

Eberhart and Kennedy [9] introduced the Particle Swarm Optimization (PSO) strategy that considers a population of candidate solutions or particles moving around the search space and are updated to their localBest or globalBest computed using their position and velocity parameters. The standard PSO algorithm suffers from major problems like the ability to explore new search spaces.

III MATERIALS AND METHODS

In this section presents the materials and methods of this research work. The dataset borrowed from <https://archive.ics.uci.edu/ml/datasets/Wireless+Indoor+Localization>. The data has collected in indoor space by observing signal strengths of seven WiFi signals visible on a smartphone. The decision variable is one of the four rooms. This data set is collected to perform experimentation on how wifi signal strengths can be used to determine one of the indoor locations. The dataset contains 8 attributes and 2000 instances. Due to limitations in this research work applies only 10% (200 instances) of instances taken from the dataset for build the proposed systems. These instances randomly taken from the dataset. Each attribute is wifi signal strength observed on smartphone. In this research work has implemented in Weka3.8.3.Version.

- The polynomial kernel : $K(x, y) = \langle x, y \rangle^p$ or $K(x, y) = (\langle x, y \rangle + 1)^p$
- NormolisedPolykernel: The normalized polynomial kernel.
- The RBF kernel : $K(x, y) = \exp(-\gamma \cdot \|x - y\|^2)$
- Puk: The Pearson VII function-based universal kernel.

IV EXPERIMENTS AND DISCUSSION

In this section discusses results and analysis of this research work. In this work applied various kernels like namely Polykernel, NormolisedPolykernel, RBFKernel and Puk and measured various novel performances like Correlation coefficient, Mean absolute error, Root mean squared error, Relative absolute error, Root relative squared error, Time Taken to build the Model(In Seconds), and Kernels Evaluations.

Table 1: Comparisons of Subset Measurement and Superset Measurement in Wireless Indoor Localization Data Set

No	Kernel	Total Number of Instances	Correlation coefficient	Mean absolute error	Root mean squared error	Relative absolute error	Root relative squared error	Time Taken to build the Model (In Seconds)	Kernels Evaluations
1	Polykernel	200	0.90	0.35	0.45	38.76 %	43.50 %	0.21	20100 (92.45% cached)
		2000	0.88	0.24	0.35	42.77%	46.99%	2.77	38870625 (23.04% cached)
2	NormolisedPolykernel	200	0.79	0.46	0.65	51.26%	62.29 %	0.11	20100 (97.15% cached)
		2000	0.88	0.25	0.43	40.41 %	47.95 %	3.28	22249277 (75.85% cached)
3	RBFKernel	200	0.84	0.56	0.66	62.94 %	63.60 %	0.18	20100 (96.68% cached)
		2000	0.88	0.27	0.41	44.13%	48.73 %	11.64	93746593 (10.60% cached)
4	Puk	200	0.98	0.13	0.23	14.52 %	22.14 %	0.18	20100 (98.49% cached)
		2000	0.98	0.05	0.12	9.51 %	17.68 %	12.25	29585338 (81.37% cached)

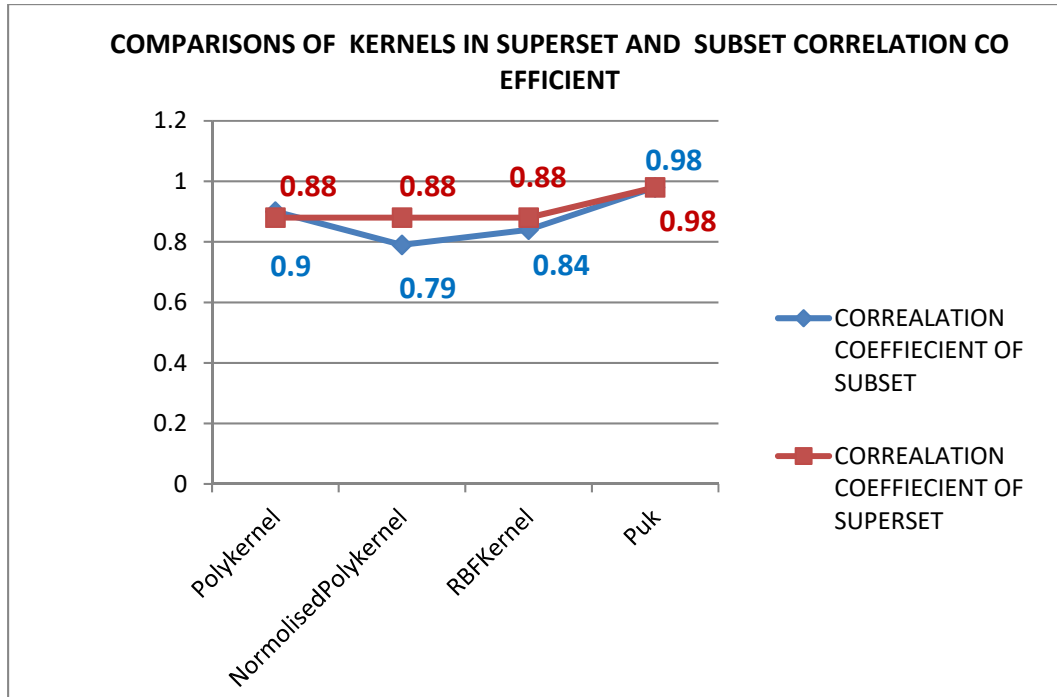


Figure 1: Various Kernels Vs Correlation coefficient of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels have various correlation co efficient values in subset of Wireless Indoor Localization Data Set. Polykernel has 0.90, NormalisedPolykernel has 0.79, RBFKernel has 0.84 and Puk Kernel has 0.98. So, Puk Kernel has highest value compare than other kernels correlation co efficient values.

Above diagram clearly depicts on various kernels have various correlation co efficient values in superset of Wireless Indoor Localization Data Set. Polykernel has 0.88, NormalisedPolykernel has 0.88, RBFKernel has 0.88 and Puk Kernel has 0.98. So, Puk Kernel has highest value compare than other kernels correlation co efficient values.

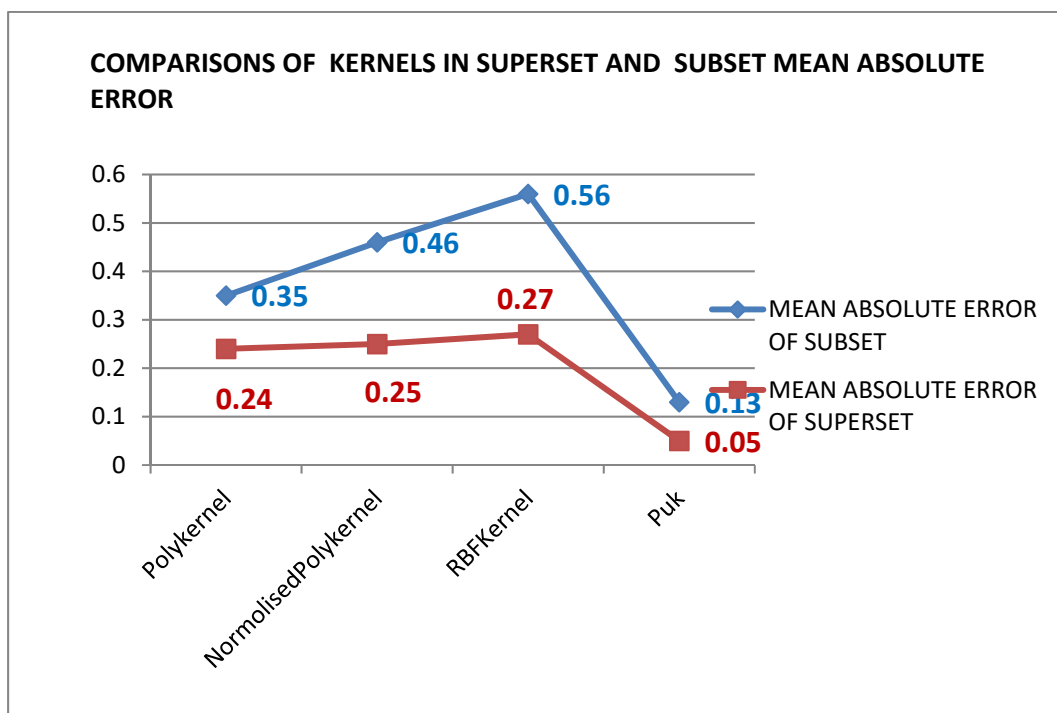


Figure 2: Various Kernels Vs Mean absolute error of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels have various Mean absolute error values in subset of Wireless Indoor Localization Data Set. Polykernel has 0.35, NormolisedPolykernel has 0.46, RBFKernel has 0.56 and Puk Kernel has 0.13. So, Puk Kernel has lowest Mean absolute error value compare than other kernels.

Above diagram clearly depicts on various kernels have various Mean absolute error values in superset of Wireless Indoor Localization Data Set. Polykernel has 0.24, NormolisedPolykernel has 0.25, RBFKernel has 0.27 and Puk Kernel has 0.05. So, Puk Kernel has lowest Mean absolute error value compare than other kernels.

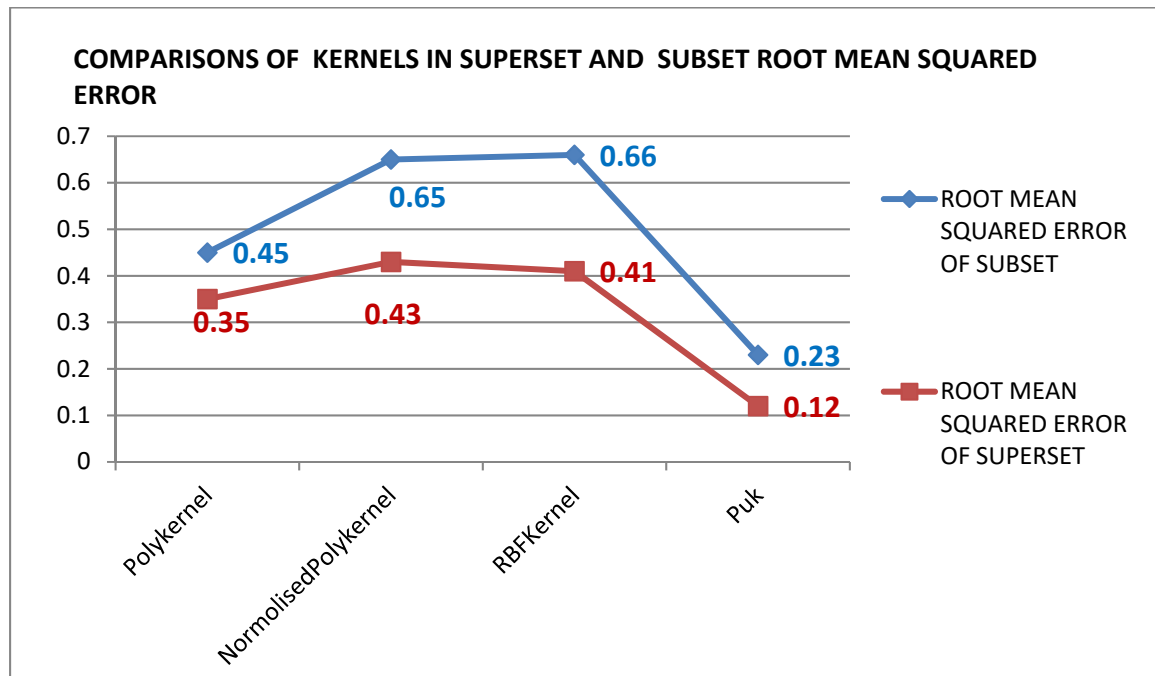


Figure 3: Various Kernels Vs Root mean squared error of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels have various Root mean squared error values in subset of Wireless Indoor Localization Data Set. Polykernel has 0.45, NormolisedPolykernel has 0.65, RBFKernel has 0.66 and Puk Kernel has 0.23. So, Puk Kernel has lowest Root mean squared error value compare than other kernels.

Above diagram clearly depicts on various kernels have various Root mean squared error values in superset of Wireless Indoor Localization Data Set. Polykernel has 0.35, NormolisedPolykernel has 0.43, RBFKernel has 0.41 and Puk Kernel has 0.12. So, Puk Kernel has lowest Root mean squared error value compare than other kernels.

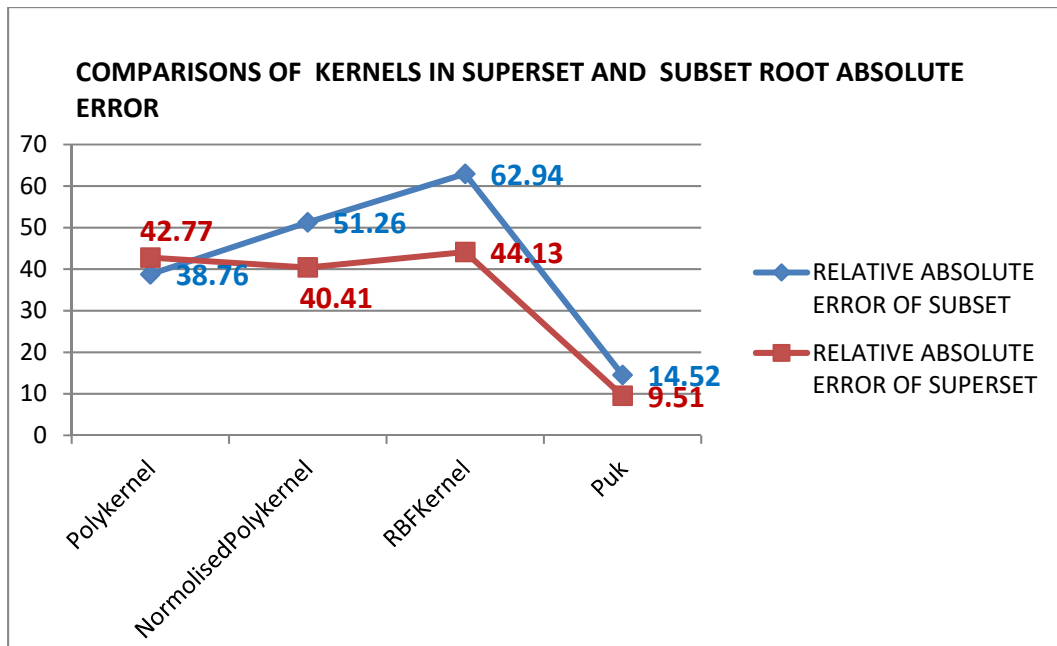


Figure 4: Various Kernels Vs Relative absolute error of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels have various Relative absolute error values in subset of Wireless Indoor Localization Data Set. Polykernel has 38.76, NormalisedPolykernel has 51.26, RBFKernel has 62.94 and Puk Kernel has 14.52. So, Puk Kernel has lowest Relative absolute error value compare than other kernels.

Above diagram clearly depicts on various kernels have various Relative absolute error values in superset of Wireless Indoor Localization Data Set. Polykernel has 42.77, NormalisedPolykernel has 40.41, RBFKernel has 44.13 and Puk Kernel has 9.51. So, Puk Kernel has lowest Relative absolute error value compare than other kernels.

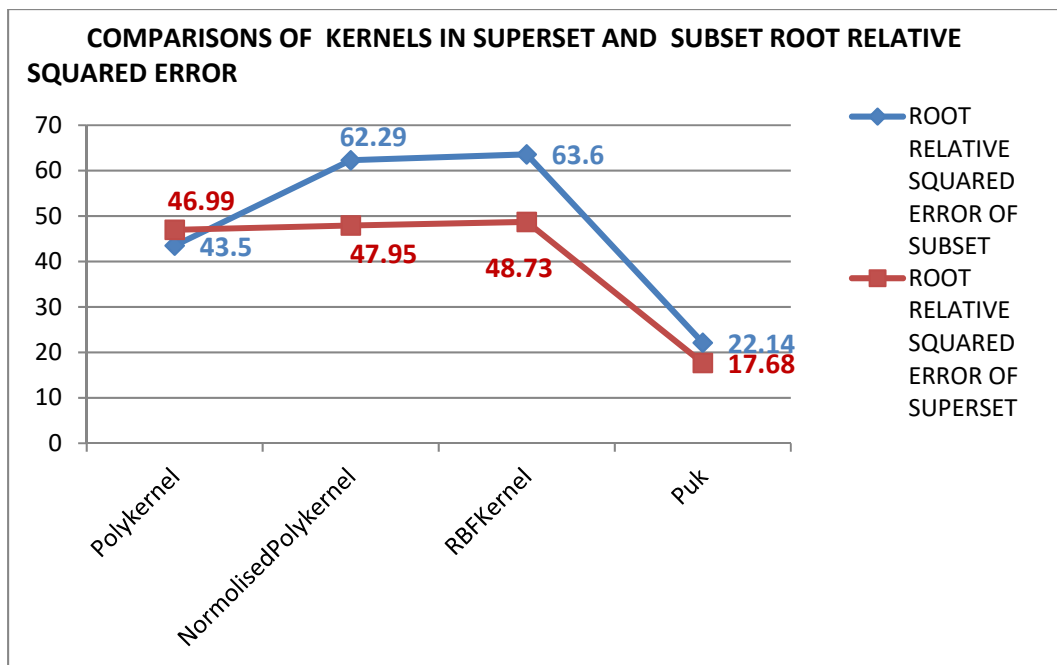


Figure 5: Various Kernels Vs Root relative squared error of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels have various Root relative squared error values in subset of Wireless Indoor Localization Data Set. Polykernel has 43.5, NormalisedPolykernel has 62.29, RBFKernel has

63.6 and Puk Kernel has 22.14. So, Puk Kernel has lowest Root relative squared error value compare than other kernels.

Above diagram clearly depicts on various kernels have various Root relative squared error values in superset of Wireless Indoor Localization Data Set. Polykernel has 46.99, NormolisedPolykernel has 47.95, RBFKernel has 48.73 and Puk Kernel has 17.68. So, Puk Kernel has lowest Root relative squared error value compare than other kernels.

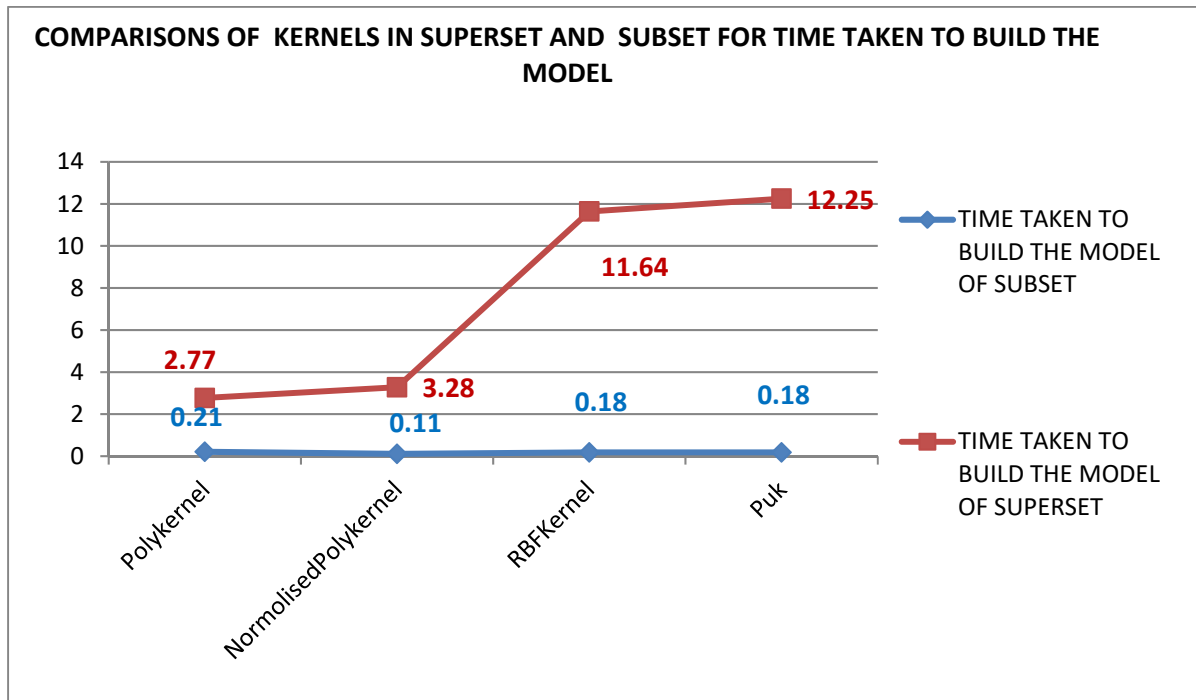


Figure 6: Various Kernels Vs Time taken to build the Model (In Seconds) of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels Vs time taken for build the model in subset of Wireless Indoor Localization Data Set. Polykernel has 0.21seconds, NormolisedPolykernel has 0.11 seconds, RBFKernel has 0.18seconds and Puk Kernel has 0.18seconds. So, NormolisedPolykernel Kernel has less time to take build the model compares than other kernels.

Above diagram clearly depicts on various kernels Vs time taken for build the model in superset of Wireless Indoor Localization Data Set. Polykernel has 2.77 seconds, NormolisedPolykernel has 3.28 seconds, RBFKernel has 11.64 seconds and Puk Kernel has 12.25 seconds. So, NormolisedPolykernel Kernel has less time to take build the model compares than other kernels.

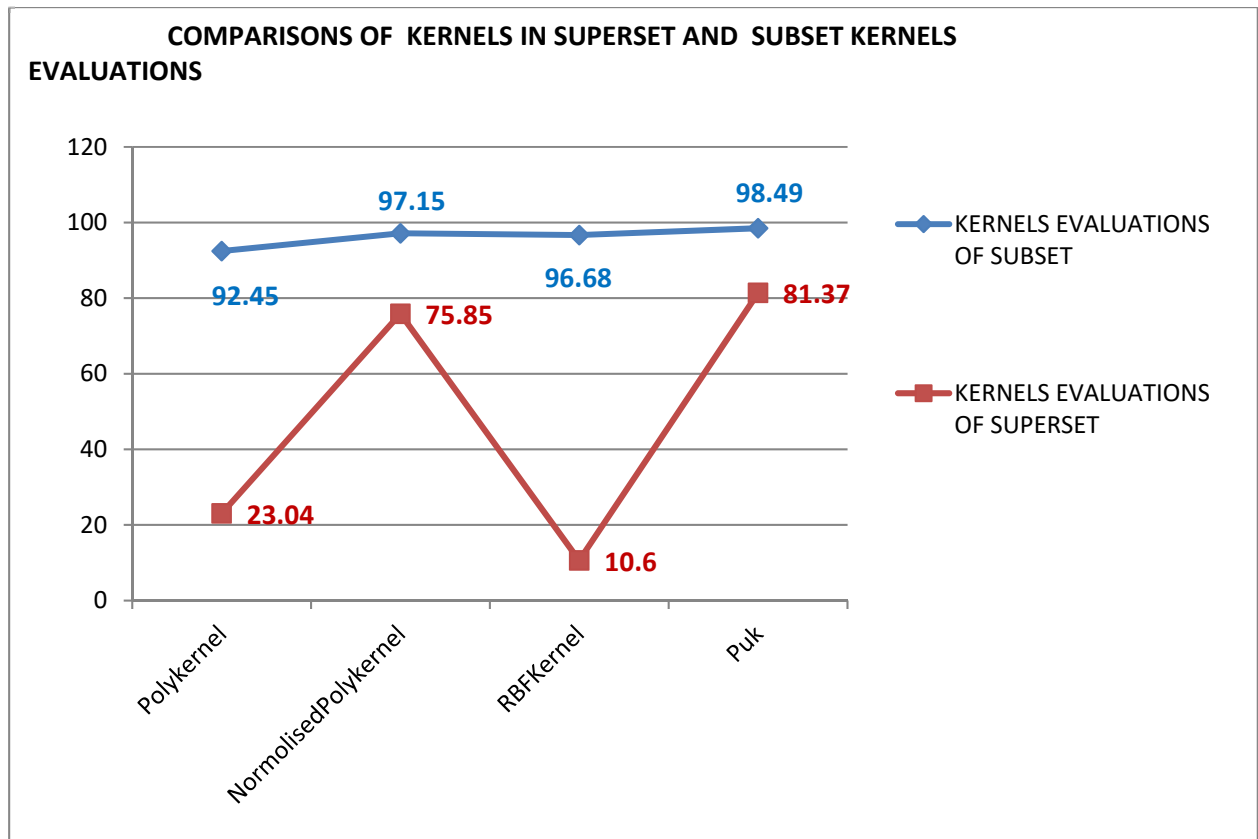


Figure 7: Various Kernels Vs Kernels Evaluations of subset and superset in Wireless Indoor Localization Data Set

Above diagram clearly depicts on various kernels Vs Kernels Evaluations in subset of Wireless Indoor Localization Data Set. Polykernel has 92.45, NormolisedPolykernel has 97.15, RBFKernel has 96.68 and Puk Kernel has 98.49. So, Puk Kernel has highly evaluated in kernels compares than other kernels.

Above diagram clearly depicts on various kernels Vs Kernels Evaluations in subset of Wireless Indoor Localization Data Set. Polykernel has 23.04, NormolisedPolykernel has 75.85, RBFKernel has 10.6 and Puk Kernel has 81.37. So, Puk Kernel has highly evaluated in kernels compares than other kernels.

V CONCLUSION

We performed tenfold cross validation and recorded the average of ten folds as the classification accuracy. From the results obtained we conclude that the support vector machine using regression with Puk kernel has less error and high co relation coefficient and it has taken less time complexity to build the model compare than other kernel parameters tuning in Support Vector Machine classifier.

REFERENCES

- [1] Jayant G. Rohra et.al, User Localization in an Indoor Environment Using Fuzzy Hybrid of Particle warm Optimization & Gravitational Search Algorithm with Neural Networks, Proceedings of Sixth International Conference on Soft Computing for Problem Solving, Advances in Intelligent Systems and Computing 546,2017.
- [2] Salazar, A.M., Warden, D.L., Schwab, K., Spe.ctor, J., Braverman, S., Walter, J.,Ellenbogen, R.G.: Cognitive rehabilitation for traumatic brain injury a randomized trial.JAMA 283(23), 3075–3081 (2000)
- [3] Nguyen, N.T., Bui, H.H., Venkatsh, S., West, G.: Recognizing and monitoring high-level behaviors in complex spatial environments. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, II-620. IEEE (2003)
- [4] Pei, L., Liu, J., Guinness, R., Chen, Y., Kuusniemi, H., Chen, R.: Using LS-SVM based motion recognition for smartphone indoor wireless positioning. Sensors 12(5), 6155–6175 (2012)
- [5] Cho, S.B.: Exploiting machine learning Techniques for location recognition and prediction with smartphone Logs. Neurocomputing (2015)
- [6] Zou, H., Lu, X., Jiang, H., Xie, L.: A fast and precise indoor localization algorithm based on an online sequential extreme learning machine. Sensors 15(1), 1804–1824 (2015)
- [7] Zadeh, L.A.: Fuzzy sets. Inform. control 8(3), 338–353 (1965)
- [8] Jang, J.S., Sun, C.T.: Neuro-fuzzy modeling and control. Proc. IEEE 83(3), 378–406 (1995)
- [9] Eberhart, R.C., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, vol. 1, pp. 39–43 (1995)
- [10] Bulusu, N., Heidemann, J., Estrin, D.: GPS-less low-cost outdoor localization for very smalldevices. IEEE Personal Commun. 7(5), 28–34 (2000).

- [11] Dr.G.Ayyappan Ensemble Classifications for Student Academics Performance Data Set Indian Journal of Computer Science and Engineering 10(1) (2019) DOI : 10.21817/indjcse/2019/v10i1/191001009
- [12] Dr.G.Ayyappan Ensemble Classifications for Sequential minimal optimization classification approach for caesarian section classification dataset data set by applying various kernels 9(6) (2018) DOI : 10.21817/indjcse/2018/v9i6/180906012