IV. CONCLUSIONS

This research has successfully accomplished the objectives where five classification techniques (Naïve Bayes, Decision Tree, Support Vector Machine, Neural Network and Random Forest) were performed for Malaysian rainfall prediction. The main objective of this study is to identify the best technique for rainfall prediction. Hence, after applying the five techniques, a comparative analysis has been performed in order to determine the most appropriate technique. The experimental results showed that for Rainfall prediction, Decision Tree, and Random Forest perform well because of their abilities to train on little data and predict the higher portion of data with higher F-measure. Support Vector Machine and Naive Bayes also trained on a small portion of data to predict higher portion but with lower F-measure. Neural Network it is an efficient method but it needs a large portion of training data to predict the very small portion of testing data. In addition, we can conclude that with small training data (10%) from 1581 instances Random Forest correctly classified 1043 instances. This result put Random Forest in the forefront of the five techniques we have been used.

For future work, the following suggestions can be considered; Combining two or more prediction algorithms has the ability to enhance the process of predicting; Use more valuable features that can generalize or discriminate the classes has a significant impact on the effectiveness; Exploit the rainfall prediction has a significant impact on predicting flowed where there is a direct correlation between the rainfalls and flowed; Use more dataset and explore more areas and locations in the world would be a valuable idea.

ACKNOWLEDGMENT

This work was supported by the Ministry of Education, Malaysia (Grant Code FRGS/2/2013/ICT02/UKM/02/).

REFERENCES

[1] A. Rajaraman and J. D. Ullman, *Mining of massive Datasets*, 20th ed. Cambridge: Cambridge University Press (Virtual Publishing), 2012.

[2] D. J. Hand, H. Mannila, P. Smyth, and D. J. H, *Principles of data mining (Adaptive computation and machine learning)*. Cambridge, MA: Bradford Books, 2001.

[3] C. L. Wu and K. W. Chau, "Prediction of rainfall time series using modular soft computing methods,"*Engineering Applications of Artificial Intelligence*, vol. 26, no. 3, pp. 997–1007, Mar. 2013.

[4] S. Rani and G. Sikka, "Recent techniques of clustering of time series data: A survey," *International Journal of Computer Applications*, vol. 52, no. 15, pp. 1–9, Aug. 2012.

[5] H. S. Badr, B. F. Zaitchik, and A. K. Dezfuli, "A tool for hierarchical climate regionalization," *Earth Science Informatics*, vol. 8, no. 4, pp. 949–958, May 2015.

[6] I. Panel, C. Change, and P. Ivonne, Climate change 2013: The physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change, Intergovernmental Panel on Climate Change, Ed. Cambridge: Cambridge University Press, 2014.

[7] K. Abhishek, A. Kumar, R. Ranjan, and S. Kumar, "A rainfall prediction model using artificial neural network," in Control and System Graduate Research Colloquium (ICSGRC), 2012, IEEE, 2012. [Online]. Available: 10.1109/ICSGRC.2012.6287140. Accessed: Nov. 6, 2016.

[8] R. Venkata Ramana, B. Krishna, S. R. Kumar, and N. G. Pandey, "Monthly rainfall prediction using Wavelet neural network analysis," *Water Resources Management*, vol. 27, no. 10, pp. 3697–3711, Jun. 2013.

[9] B. Wang et al., "Rethinking Indian monsoon rainfall prediction in the context of recent global warming," Nature Communications, vol. 6, p. 7154, May 2015.

[10] K. K. Htike and O. O. Khalifa, "Rainfall forecasting models using focused time-delay neural networks,"*International Conference on Computer and Communication Engineering (ICCCE'10)*, May 2010.

[11]D. Gupta and U. Ghose, "A comparative study of classification algorithms for forecasting rainfall," *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)*, Sep. 2015.

[12] J. Joseph and R. T K, "Rainfall prediction using data mining techniques," *International Journal of Computer Applications*, vol. 83, no. 8, pp. 11–15, Dec. 2013.

[13] "The Oryx resource guide to El Nino and la Nina," *Choice Reviews Online*, vol. 40, no. 04, pp. 40–2226–40–2226, Dec. 2002.

[14] J. Kajornrit, K. W. Wong, and C. C. Fung, "Rainfall prediction in the northeast region of Thailand using modular fuzzy inference system," *2012 IEEE International Conference on Fuzzy Systems*, Jun. 2012.

[15] V. B. Nikam and B. B. Meshram, "Modeling rainfall prediction using data mining method: A Bayesian approach," *2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation*, Sep. 2013.

[16] M. Kühnlein, T. Appelhans, B. Thies, and T. Nauss, "Improving the accuracy of rainfall rates from optical satellite sensors with machine learning — A random forests-based approach applied to MSG SEVIRI,"*Remote Sensing of Environment*, vol. 141, pp. 129–143, Feb. 2014.

[17] D. Isa, L. H. Lee, V. P. Kallimani, and R. RajKumar, "Text document Preprocessing with the Bayes formula for classification using the support vector machine," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 9, pp. 1264–1272, Sep. 2008.

[18] R. S. V. Teegavarapu and V. Chandramouli, "Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records," *Journal of Hydrology*, vol. 312, no. 1-4, pp. 191–206, Oct. 2005.

[19] S. S. Monira, Z. M. Faisal, and H. Hirose, "Comparison of artificially intelligent methods in short term rainfall forecast," *2010 13th International Conference on Computer and Information Technology (ICCIT)*, Dec. 2010.