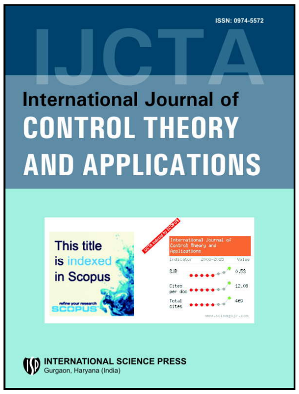

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 14 • 2017

ISPM-OC: Improved Snow Prediction Model Using Optimal k-Means Clustering and Decision Tree to Nowcast Snow/No-Snow

C. Kishor Kumar Reddy[1] and B. Vijaya Babu[2]

[1] *Research Scholar, Dept. of CSE K L University, Guntur, AP, India, Email: Kishoar23@gmail.com*
[2] *Professor, Dept. of CSE K L University, Guntur, AP, India, Email: Vijaymtech28@gmail.com*

Abstract: Long before technology was developed, folks had to trust patterns, observations and their expertise to Nowcast snow/no-snow. In this paper, a new algorithm is proposed, based on the concepts of clustering and decision tree approaches using historical weather datasets. The algorithm does not use the conventional decision tree approach for identifying the split points; instead it introduces clustering mechanism to select split points. Concepts of clustering is used to find the split points, concepts of decision tress is used to find the best split point, in which entropy is opted as attribute selection measure. This offers better opportunity for data mining, and inherently provides an effective method for nowcasting snow/no-snow.

Keywords: Clustering, Decision Trees, Entropy, Nowcasting, Snow

1. INTRODUCTION

Nowcasting combines a description of the current state of the atmosphere and a short-term forecast of how the atmosphere will evolve during the next several hours. A convergence of technical developments has set the stage for a major jump in nowcasting capabilities and the ability to apply those advances to important societal needs. Accurate and timely nowcasting of snow/no-snow is a major challenge for the scientific community. Snowfall nowcasting modeling involves a combination of computer models, observation and knowledge of trends and patterns. Using these methods, reasonably accurate forecasts can be made up. Several recent research studies have developed snowfall nowcasting using different weather and climate forecasting methods [8] [17] [24] [26-35] [48-49]. For detailed literature, refer our previous papers [45-47].

In this paper, we propose Improved Snow Prediction Model using Optimal k-means Clustering (ISPM-OC), which is based on an un-supervised learning method called clustering mechanism and supervised learning method called decision tree construction. In the present research, k-means clustering mechanism is integrated with SLIQ decision tree algorithm to nowcast snow/no-snow effectively. In SLIQ, at every node data is to be sorted, splits are to be identified whenever there is a change in the class label. This increases the computation of number of splits, which in-turn increases computational complexity. Hence, the proposed integrated approach ISPM-OC employs a scheme that does away with the need to sort the data at every node of the decision tree.

Instead, the training data need to be partitioned using k-means clustering only once for each numeric attribute at the beginning of the tree growth phase. In addition, the split point value is computed at the cluster boundaries at both the beginning and end of the cluster segments. Consequently, splits of all the leaves of the current tree are simultaneously adopted in one pass over the data. The new technique, called ISPM-OC, is quite different from existing methods, and it has many distinctive advantages.

The main contribution of this paper is that it proposes Improved Snow Prediction Model using optimal k-means clustering, which is based on k-means clustering mechanism and SLIQ decision tree methods. It is fundamentally different from existing decision tree techniques. Existing techniques evaluates split points whenever there is a change in the class label. The proposed technique, however, finds split points based on cluster boundaries.

a. The model is capable to nowcast snow/no-snow based on the weather attributes: humidity, temperature, pressure, wind speed, dew point and visibility more effectively.

b. The proposed model has the capability to predict weather before 4 hours more effectively.

c. A detailed evaluation against other prediction decision tree and non-decision tree algorithms is performed, that provide a fair comparison to show the effectiveness of the proposed model.

d. The proposed model is evaluated with various performance measures such as accuracy, specificity, sensitivity, precision, error rate and also in terms of number of split points.

The rest of the paper develops the idea further. Section 2 provides the description and working of the new model. Section 3 analyzes and discusses the result and finally section 4 concludes the paper with future directions and references.

2. ISPM-OC DECISION TREE ALGORITHM

The experimental implementation methodology of ISPM-OC algorithm consists of four stages: 1) a k-means algorithm to group N data points into "k" disjoint clusters, where "k" is determined by an auto detection cluster classifier algorithm explained later in this section; 2) identification of the split points; 3) evaluation of the entropy for all the attributes; and 4) decision tree construction.

Algorithm

1. Read dataset to select the root node of the ISPM-OC decision tree.
2. Generate an attribute list for each attribute of the dataset.
3. Compute the Class Entropy for each class label

$$\text{Class Entropy} = -\sum_{i=1}^{M} P_i \log_2 P_i \tag{1}$$

4. Partition the training data along with the class label on each attribute "v_q" using k-means clustering and mark the beginning and ending value positions of each cluster segments as "s_p".
5. Create two subsets for each "s_p" such that subset S_1 has values less than "s_p" and subset S_2 has values greater than or equal to "s_p".
6. Compute Attribute Entropy for each and every attribute "v_q"

$$\text{Attribute Entropy} = -\sum_{j=1}^{N} P_j \left[-\sum_{i=1}^{M} P_i \log_2 P_i \right] \tag{2}$$

7. Compute Entropy for each and every attribute "v_q"

$$Entropy = Class\ Entropy - Attribute\ Entropy \tag{3}$$

8. The maximum Entropy is considered to be the best split point and becomes the root node.
9. Repeat Steps 6 through 8, generating leaf nodes in place of the root node until all leaf nodes contain the same class labels.

A fundamental problem in k-means clustering is to determine the number of clusters, which is usually taken as prior or fixed. The selection of a good value for "k" can affect the overall accuracy of the algorithm, and clustering solutions may vary as different numbers of clusters are specified. A clustering technique would most possibly recover the underlying cluster structure, given a good estimate of the true number of clusters. To overcome the scenario, in this paper, an Optimal Decision Cluster Classifier is proposed. Choosing a value for "k" by visual inspection can be automated by using the percentage of variance of clusters that determines the optimum number of clusters. This method finds the optimal number of clusters automatically, based on the relationship between consecutive differences among the data points.

Optimal Clustering

1. Read all the records of an attribute.
2. Compute consecutive differences for all the records.
3. Repeat Step 2, till it ends with a single record value for a particular attribute.
4. Traverse from bottom to top to identify the maximum single digit value i.e., 1–9.
5. The iteration that has the maximum single digit value is taken to be the optimal cluster size.

3. RESULTS AND DISCUSSION

The performance of the ISPM-OC algorithm was compared with the other decision tree algorithms: Decision Stump, J48, LMT, Random Forest, REP Tree, SLIQ, SPM, SLGAS, ISLIQ, ISPM, ISLGAS and ISLIQ-OC and non-decision tree algorithms: Bayes Net, Naïve Bayes, Multilayer Perceptron, SMO and Simple Logistic using 20 international locations snow/no-snow datasets taken from the www.wunderground.com. All results that we subsequently report are based on tenfold cross validation. Table 1 provides a detailed illustration on the dataset.

The comparison of split points is discussed in Table 2, number of split points are significantly better (shown in boldface font) for most of the datasets. The comparison of classification accuracy with our previous developed decision tree algorithms is discussed in Table 3. The classification accuracy is significantly better for most of the datasets, except for the few datasets where the accuracy is marginally less. However, on an average the proposed method yields an average accuracy of 89.64%. The comparison of classification accuracy with non-decision tree algorithms is discussed in Table 4. The classification accuracy is significantly better for most of the datasets, except for the few datasets where the accuracy is marginally less. However, on an average the proposed method yields an average accuracy of 89.64%. The comparison of classification accuracy with existing decision tree algorithms is discussed in Table 5. The classification accuracy is significantly better for most of the datasets, except for the few datasets where the accuracy is marginally less. However, on an average the proposed method yields an average accuracy of 89.64%.

The comparison of classification error-rate with other decision tree algorithms is discussed in Table 6. The classification error-rate is significantly better for most of the datasets, except for the few datasets where the error-rate is marginally more. However, on an average the proposed method yields an average error-rate of 10.35%.

Table 1
Data set Description

City Name	*Instances*	*Training*	*Testing*	*Attributes*	*Classes*
Aberdeen	6333	4750	1583	5	2
Bangkok	5740	4305	1435	5	2
Barcelona	6013	4510	1504	5	2
Benton	23042	17281	5761	5	2
Botswana	6047	4535	1512	5	2
Brazil	6367	4775	1592	5	2
Cairo	6143	4607	1536	5	2
Chennai	6033	4525	1508	5	2
Delhi	6015	4511	1504	5	2
Eglinton	6318	4738	1580	5	2
Humberside	1036	777	259	5	2
Hyderabad	5849	4387	1462	5	2
Iceland	3512	2634	878	5	2
Lahore	4887	3665	1222	5	2
Manchester	6338	4753	1585	5	2
Norway	6105	4579	1526	5	2
Perth	6182	4636	1546	5	2
Sellaness	5412	4059	1353	5	2
Tiruptahi	6039	4529	1510	5	2
Valley	6082	4561	1521	5	2

Further, the size of the decision tree constructed, time taken and number of rules computed using the proposed algorithm is significantly less compared with that of other decision tree algorithms.

Table 2
Comparison of Split Points

City Name	*SPM*	*ISPM*	*ISPM-OC*
Aberden	740	210	**65**
Bangkok	112	50	75
Barcelona	174	41	67
Benton	449	208	**81**
Botswana	195	88	**67**
Brazil	460	299	**79**
Cairo	165	162	**67**
Chennai	130	64	69
Delhi	281	162	**66**
Eglinton	360	43	**69**
Humberside	171	34	**58**
Hyderabad	116	72	**68**
Iceland	385	207	**74**
Lahore	190	51	**69**
Manchester	499	211	**227**
Norway	765	560	**67**
Perth	246	136	**67**
Sellaness	391	154	**65**
Tirupathi	154	108	**71**
Valley	706	200	854
Average	**334.5**	**153**	**116.25**

Table 3
Comparison of accuracy with previous decision tree algorithms

City Name	*SLIQ*	*SPM*	*SLGAS*	*ISLIQ*	*ISPM*	*ISLGAS*	*ISLIQ-OC*	*ISPM-OC*
Aberden	87.3	85.47	85.97	87.61	87.68	88.98	85.84	85.844
Bangkok	96.09	94.49	95.19	98.11	98.32	97.9	98.53	96.3
Barcelona	95.8	95.14	95.67	96.07	96.07	96.07	95	95.27
Benton	70.05	70.14	72.03	70.12	70.24	70.41	69.93	68.14
Botswana	93.78	96.16	93.58	95.43	96.29	96.62	98.21	98.14
Brazil	75.5	73.05	75.75	73.36	75.18	75.6	71.98	75.37
Cairo	88.99	89.7	89.77	89.98	89.32	90.1	91.6	91.53
Chennai	76.65	76.12	77.51	76.35	74.6	72.08	82.82	81.43
Delhi	96.14	94.94	96.8	93.15	96.34	96.8	95.74	92.95
Eglinton	89.24	90.06	90.06	89.56	89.75	89.87	89.24	89.37
Humberside	93.05	94.59	94.98	93.82	94.2	94.82	94.2	93.05
Hyderabad	96.5	97.8	94.79	96.4	97.8	97.67	97.6	98.15
Iceland	89.17	88.49	90.2	88.49	88.95	87.81	87.47	88.61
Lahore	84.82	86.05	85.89	84.65	86.38	85.06	84.65	84.74
Manchester	92.74	92.87	89.58	93.43	91.29	92.36	92.11	92.49
Norway	88.99	90.89	90.62	90.89	90.69	90.3	89.18	90.82
Perth	94.3	94.43	96.31	94.43	94.24	94.37	94.37	95.14
Sellaness	75.9	77.67	79.45	84.4	84.18	84.7	83	83.66
Tirupathi	97.54	97.41	97.41	97.48	97.54	97.35	98.87	99
Valley	90	91.38	90.32	91.3	90.52	91.76	92.36	92.96
Average	88.62	88.84	89.09	89.25	89.47	89.53	89.63	89.64

Table 4
Comparison of accuracy with non-decision tree algorithms

City Name	*Bayes Net*	*Naïve Bayes*	*Multilayer Perceptron*	*SMO*	*Simple Logistic*	*ISPM-OC*
Aberdeen	82.53	80.66	85.86	80.82	81.47	85.844
Bangkok	98.95	97.91	98.95	98.95	98.95	96.3
Barcelona	96.27	96.93	98.6	98.33	98.4	95.27
Benton	65.55	63.83	68.35	65.36	65.41	68.14
Botswana	99.27	99.4	99.4	99.2	99.13	98.14
Brazil	74.41	74.41	76.86	74.03	74.41	75.37
Cairo	95.83	96.67	97.65	97.52	97.65	91.53
Chennai	87.73	82.09	85.95	85.95	85.95	81.43
Delhi	86.16	85.29	90.48	89.28	88.95	92.95
Eglinton	94.49	95.69	95.75	95.18	95.69	89.37
Humberside	84.32	83.14	84.18	85.33	84.71	93.05
Hyderabad	99.31	99.31	99.31	99.31	99.31	98.15
Iceland	85.64	82	86.33	83.48	82.34	88.61
Lahore	83.85	80.09	86.96	87.21	86.97	84.74
Manchester	86.04	87.05	89.14	87.24	87.75	92.49
Norway	86.15	85.14	84.33	86.79	84.31	90.82
Perth	88.18	87.15	88.19	87.44	87.39	95.14
Sellaness	88.76	86.17	86.5	87.58	88.17	83.66
Tiruptahi	88.53	84.36	95.69	95.56	95.56	99
Valley	88.18	87.17	89.53	88.75	89.14	92.96
Average	88	86.72	89.4	88.66	88.58	89.64

Table 5
Comparison of accuracy with existing decision tree algorithms

City Name	*Decision*	*J48*	*LMT*	*Random Forest*	*REP Tree*	*ISPM-OC*
Aberdeen	78.9	89.19	86.44	99.03	87.98	85.84
Bangkok	98.95	98.95	98.95	98.74	98.95	96.3
Barcelona	98.33	98.33	98.4	98.33	98.33	95.27
Benton	63.29	68.48	68.52	66.18	68.24	68.14
Botswana	99.2	99.27	99.13	99.4	99.2	98.14
Brazil	68.44	65.85	67.24	67.99	65.92	75.37
Cairo	97.52	97.85	97.65	97.39	97.78	91.53
Chennai	85.95	85.95	85.95	85.35	86.02	81.43
Delhi	80.5	89.85	88.95	88.02	89.95	92.95
Eglinton	85.18	86.45	86.26	86.07	86.2	89.37
Humberside	91.35	91.28	91.65	92.32	91.55	93.05
Hyderabad	89.31	89.31	89.31	89.31	89.31	98.15
Iceland	81.77	87.35	87.47	85.53	86.44	88.61
Lahore	84.84	87.05	86.96	86.47	87.05	84.74
Manchester	84.72	88.63	88.06	87.24	88	92.49
Norway	88.41	86.49	86.11	86.31	86.69	90.82
Perth	91.42	93.21	90.73	92.69	91.76	95.14
Sellaness	78.83	80.24	81.27	80.09	79.87	83.66
Tiruptahi	95.56	95.56	95.56	95.42	95.56	99
Valley	88.75	89.14	89.73	89.01	89.07	92.96
Average	85.56	88.42	88.21	88.54	88.19	89.64

Table 6
Comparison of error rate with other decision tree algorithms

City Name	*SLIQ*	*SPM*	*SLGAS*	*ISLIQ*	*ISPM*	*ISLGAS*	*ISLIQ-OC*	*ISPM-OC*
Aberdeen	12.7	14.53	14.03	12.39	12.32	11.02	14.16	14.156
Bangkok	3.91	5.51	4.81	1.89	1.68	2.1	1.47	3.7
Barcelona	4.2	4.86	4.33	3.93	3.93	3.93	5	4.73
Benton	29.95	29.86	27.97	29.88	29.76	29.59	30.07	31.86
Botswana	6.22	3.84	6.42	4.57	3.71	3.38	1.79	1.86
Brazil	24.5	26.95	24.25	26.64	24.82	24.4	28.02	24.63
Cairo	11.01	10.3	10.23	10.02	10.68	9.9	8.4	8.47
Chennai	23.35	23.88	22.49	23.65	25.4	27.92	17.18	18.57
Delhi	3.86	5.06	3.2	6.85	3.66	3.2	4.26	7.05
Eglinton	10.76	9.94	9.94	10.44	10.25	10.13	10.76	10.63
Humberside	6.95	5.41	5.02	6.18	5.8	5.18	5.8	6.95
Hyderabad	3.5	2.2	5.21	3.6	2.2	2.33	2.4	1.85
Iceland	10.83	11.51	9.8	11.51	11.05	12.19	12.53	11.39
Lahore	15.18	13.95	14.11	15.35	13.62	14.94	15.35	15.26
Manchester	7.26	7.13	10.42	6.57	8.71	7.64	7.89	7.51
Norway	11.01	9.11	9.38	9.11	9.31	9.7	10.82	9.18
Perth	5.7	5.57	3.69	5.57	5.76	5.63	5.63	4.86
Sellaness	24.1	22.33	20.55	15.6	15.82	15.3	17	16.34
Tiruptahi	2.46	2.59	2.59	2.52	2.46	2.65	1.13	1
Valley	10	8.62	9.68	8.7	9.48	8.24	7.64	7.04
Average	11.4	11.2	10.9	10.74	10.5	10.46	10.36	10.35

4. CONCLUSION

The IPSM-OC algorithm has outperformed when compared with the other decision tree algorithms: Decision Stump, J48, LMT, Random Forest, REP Tree, SLIQ, SPM, SLGAS, ISLIQ, ISPM, ISLGAS and ISLIQ-OC and non-decision tree algorithms: Bayesian Networks, Naïve Bayes, Multilayer Perceptron, SMO and Simple Logistic over 20 international locations snow/no-snow datasets taken from the www.wunderground.com.The classification accuracy, sensitivity, specificity, error rate and number of rules are significantly better in the case of proposed algorithm compared to that of previous algorithms. Huge reduction in number of split points during the construction of the decision tree over the majority datasets is, on average, observed for the ISPM-OC algorithm in comparison to other decision tree algorithms.

REFERENCES

[1] Robert A. Houze: Cloud Dynamics. Academic Press (1994).

[2] Nabilah Filzah Mohd Radzuan, Andi Putra, Zalinda Othman, Azuraliza Abu Bakar and Abdul Razak Hamdan: Comparative Study – Three Artificial Intelligence Techniques for Rain Domain in Precipitation Forecast. International Journal of Environmental, Ecological, Geological and Mining Engineering (2013) 898-903.

[3] http://en.wikipedia.org/wiki/Weather_forecasting.

[4] http://en.wikipedia.org/wiki/Quantitative_precipitation_forecast.

[5] J. Han: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers. (2001)

[6] Irene Y.H. Gu, Unai Sistiag and Anders Fahlstrom: Online Detection of Snow Coverage and Swing Angles of Electrical Insulators on Power Transmission Lines. IEEE International Conference on Image Processing (2009) 3249 – 3252.

[7] Jinmei Pan, Lingmei Jiang and Lixin Zhang: Wet Snow Detection in the South of China by Passive Microwave Remote Sensing. IEEE International Geo Science and Remote Sensing Symposium (2012) 4863-4866.

[8] Yajaira Mejia, Hosni Ghedira and Reza Khanbilvardi: A Neural Network Based Approach for Multi-Spectral Snowfall Detection and Estimation, IEEE International Geo Science and Remote Sensing Symposium (2007) 2276 – 2279.

[9] Melanie Wetzel, Michael Meyers, Randolph Borys, Ray Mcanelly, William Cotton, Andrew Rossi, Paul Frisbie, David Nadler, Douglas Lowenthal, Stephen Cohn, and William Brown: Mesoscale Snowfall Prediction And Verification In Mountainous Terrain. AMS Journal Of Weather And Forecasting (2004) 806-828.

[10] Pascal Sirguey, Renaud Mathieu Yves Arnaud, Muhammad M. Khan and Jocelyn Chanussot: Improved Resolution For The Detection Of Snow With Modis Using Wavelet Fusion. IEEE International Geo Science and Remote Sensing Symposium (2007) 3975-3978.

[11] Michael A. Rotondi: Estimating Transition Matrices to Predict Tomorrow's Snowfall Using Real Data. Journal of Statistics Education (2010) 1-14.

[12] Gail M. Sko Fronick Jackson, Benjamin T. Johnson, and S. Joseph Munchak: Detection Thresholds of Falling Snow From Satellite-Borne Active And Passive Sensors, IEEE Transactions On Geo Science and Remote Sensing (2013) 4177-4189.

[13] Gail M. Skofronick Jackson and Benjamin T. Johnson: Thresholds Of Detection For Falling Snow From Satellite-Borne Active And Passive Sensors, IEEE International Geo Science And Remote Sensing Symposium (2007) 2637-2640.

[14] Andrea Spisni, Fausto Tomei, Sara Pignone, Enrico Muzzi, Alessandro Panzacchi, Gabriele Antolini, Giulia Villani, Michele Di Lorenzo, Rosanna Foraci1, Marco Bittelli and Erin S. Brooks: Snow Cover Analysis in Emilia-Romagna. European Journal of Remote Sensing (2011) 59-73.

[15] Alberto Martinez Vazquez and Joaquim Fortuny Guasch: Snow Avalanche Detection and Classification Algorithm for GB-SAR Imagery, IEEE International Geo Science and Remote Sensing Symposium (2007) 3740-3743.

[16] Jeremie Bossu, Nicolas Hautière and Jean Philippe Tarel: Rain or Snow Detection in Image Sequences through Use of a Histogram of Orientation of Streaks, International Journal of Computer Vision (2011) 348-367.

[17] Noel Dacruz Evora, Ddominique Tapsoba and Danielle De Seve: Combining Artificial Neural Network Models, Geo statistics,

and Passive Microwave Data for Snow Water Equivalent Retrieval and Mapping. IEEE Transactions on Geo Science and Remote Sensing (2008) 1925-1939.

[18] Hossein Zeinivand and Florimond De SmISLIQ: Simulation of Snow Covers area by a Physical Based Model. World Academy of Science, Engineering and Technology (2009) 428-433.

[19] Xiaolan Xu, Ding Liang, Leung Tsang, Konstantinos M. Andreadis, Edward G. Josberger, Dennis P. Lettenmaier, Donald W. Cline and Simon H. Yueh: Active Remote Sensing of Snow Using NMM3D/DMRT and Comparison With Clpx Ii Airborne Data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2010) 689-697.

[20] B. B. Fitzharris and B. P. Mcalevey: Remote Sensing of Seasonal Snow Cover in the Mountains of New Zealand Using Satellite Imagery. Taylor and Francis Geocarto International (1999) 35-44.

[21] Ashok N. Srivastava and Julienne Stroeve: Onboard Detection of Snow, Ice, Clouds and Other Geophysical Processes Using Kernel Methods. Proceedings of the ICML 2003 Workshop on Machine Learning Technologies for Autonomous Space Applications (2003) 1-5.

[22] G. Singh, Y. Yamaguchi, S. E. Park and G. Venkataraman: Identification of Snow Using SAR Polarimetry Techniques. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science (2010) 146-149.

[23] Fan Ke and Tian Baoqiang: Prediction of Wintertime Heavy Snow Activity in North East China. Springer Chinese Science Bulletin (2013) 1420-1426.

[24] Folorunsho Olaiya: Application of Data Mining Techniques in Weather Prediction and Climate Change Studies, International Journal of Information Engineering and Electronic Business (2012) 51-59.

[25] Manjeet Singh, V. D. Mishra, N. K. Hakur and Jyoti Dhar Sharma: Remote Sensing GIS Based Statistical Modelling for the Prediction of Natural Hazards. International Journal of Engineering Research and Technology, (2012) 1-7.

[26] J.R. Quinlan: Induction of Decision Trees. Journal of Machine Learning (1986) 81-106.

[27] B. Chandra and P. Paul Varghese: Fuzzy Sliq Decision Tree Algorithm: IEEE Transactions on Systems, Man and Cybernetics (2008) 1294-1301.

[28] Anuja Priyama, Abhijeeta, Rahul Guptaa, Anju Ratheeb and Saurabh Srivastavab: Comparative Analysis of Decision Tree Classification Algorithms, International Journal of Current Engineering and Technology (2013) 334-337.

[29] Masud Karim and Rashedur M. Rahman: Decision Tree and Naive Bayes Algorithm for Classification and Generation of Actionable Knowledge for Direct Marketing. Journal of Software Engineering and Applications (2013) 196-206.

[30] Pedro Domingos and Michael J. Pazzani: Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier. International Conference on Machine Learning (1996) 105-112.

[31] Manish Mehta, Rakesh Agarwal and Jorma Rissanen: SLIQ: A Fast Scalable Classifier for Data Mining. International Conference on Extending Database Technology (1996) 18-32.

[32] Rodrigo Coelho Barros, Marcio Porto Basgalupp, Andre C.P.L.F. De Carvalho and Alex A. Freitas: A Survey of Evolutionary Algorithms for Decision Tree Induction, IEEE Transactions on Systems, Man and Cybernetics (2012) 291-312.

[33] J.R. Quinlan: C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, (1993).

[34] S. Safavian and D. Landgrebe: A Survey of Decision Tree Classifier Methodology, IEEE Transactions on Systems, Man And Cybernetics (1991) 660 -674.

[35] Arun k pujari: Data Mining Techniques. Universities Press (2004).

[36] http://www.wunderground.com/.

[37] Powers and M. W. David: Evolution: From Precision, Recall and F-Measure to Roc, Informedness, Markedness and Correlation. Journal of Machine Learning Technologies (2011) 37-63.

[38] Kishor Kumar Reddy C, Vijaya Babu B, Rupa C H: SLEAS: Supervised Learning using Entropy as Attribute Selection Measure. International Journal of Engineering and Technology (2014) 2053-2060.

[39] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu B: A Pragmatic Methodology to Predict the Presence of Snow/No-Snow using Supervised Learning Methodologies. International Journal of Applied Engineering Research (2014) 11381-11394.

[40] Chandra, B., Mazumdar, S., Vincent, A. and Parimi, N: Elegant Decision Tree Algorithm for Classification in Data Mining. Proceedings of the Third International Conference on Web Information Systems Engineering (2002) 160"169.

[41] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu: SPM: A Fast and Scalable Model for Predicting Snow/No-Snow. World Applied Sciences Journal (2014) 1561-1570.

[42] Pramote Luenam, Supawadee Ingsriswang, Lily Ingsrisawang, Prasert Aungsuratana, and Warawut Khantiyanan: A Neuro-Fuzzy Approach for Daily Rainfall Prediction over the Central Region of Thailand. International Multi Conference of Engineers and Computer Scientists (2010).

[43] Lily Ingsrisawang, Supawadee Ingsriswang, Pramote Luenam, Premjai Trisaranuwatana, Song Klinpratoom, Prasert Aungsuratana, and Warawut Khantiyanan: Applications of Statistical Methods for Rainfall Prediction over the Eastern Thailand. International Multi Conference of Engineers and Computer Scientists (2010).

[44] Thair Nu Phyu: Survey of Classification Techniques in Data Mining. International Multi Conference of Engineers and Computer Scientists (2009).

[45] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu: SLGAS: Supervised Learning using Gain Ratio as Attribute Selection Measure to Nowcast Snow/No-Snow. International Review on Computers and Software (2015).

[46] Kishor Kumar Reddy C, Rupa C H and Vijaya Babu: ISLIQ: Improved Supervised Learning in Quest to Nowcast Snow/No-Snow. WSEAS Transactions on Computers (2015).

[47] Kishor Kumar Reddy C, Vijaya Babu: ISPM: Improved Snow Prediction Model to Nowcast Snow/No-Snow. International Review on Computers and Software (2015).

[48] Joel T. de Castro, Gabriel M. Salistre, Jr, Young-Cheol Byun and Bobby D. Gerardo," Flash Flood Prediction Model based on Multiple Regression Analysis for Decision Support System", IAENG Proceedings of the World Congress on Engineering and Computer Science (2013).

[49] Rosmina Bustami, Nabil Bessaih, Charles Bong, Suhaila Suhaili, "Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River", IAENG International Journal of Computer Science (2007).