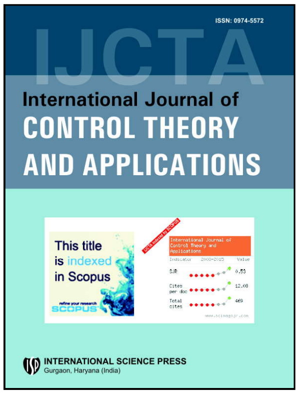

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 14 • 2017

ISLIQ-OC: Improved Supervised Learning in Quest using Optimal *k*-means Clustering Mechanism 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, Email: Kishoar23@gmail.com*
[2] *Professor, Dept. of CSE K L University, Guntur, AP, Email: Vijaymtech28@gmail.com*

Abstract: Use of conventional supervised learning mechanisms in nowcasting of snow/no-snow computes large number of split points, which in-turn increases the computational complexity and increases number of rules. This paper presents a integrated approach i.e. supervised learning algorithm SLIQ combines with clustering algorithm *k*-means, that works on minimizing split points, reduces number of rules, reduces computational complexity and increases the nowcasting rate in terms of performance. We adopted *k*-means algorithm, in order to identify the split points, in SLIQ Gini Index is adopted as attribute selection measure in order to identify the best split point. The performance of this algorithm has been studied on 20 international locations of snow/no-snow data sets. Comparisons with other existing decision/non-decision algorithms illustrate the effectiveness of this approach.

Keywords: Atmosphere, Classification, Clustering, Decision Trees, Forecasting, Gain Ratio, Snow

1. INTRODUCTION

The present research mainly focuses on nowcasting snow/no-snow. Snow disasters cause immeasurable losses to human society each year, threatening people's lives and properties and therefore attaching much importance to the nowcasting of snow/no-snow. Snowfall, however, is one of the most difficult to be measured among meteorological elements. It mainly affects transport sector such as road, air and rail ways. To solve this problem, the real-time nowcasting of snow more accurately is to be developed [5]. This can effectively avoid railway, road and air accidents caused by snowstorms, improving transport safety as well as providing a quantitative reference for the safe operation [6]. Technological improvements in the computational power are still not sufficient to handle the nowcasting efficiently. Although the current abilities of computer systems helped the meteorologist to implement more advance model that requires high computation and improves the prediction capabilities; the accuracy and timely prediction of weather phenomena is still a major issue. Further, the global climate changes and incident of some disastrous weather events increased the importance of timely and accurate weather prediction [1].

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 Improved Supervised Learning in Quest using Optimal k-means Clustering (ISLIQ-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 present research is the enhancement to our research SLIQ [31] [39], ISLIQ [46], SLEAS [38], SPM [41], ISPM [47], SLGAS [45].

Our specific contributions in this paper are listed below:

a. The major contribution of the present research is reduction of split points during decision tree growth.

b. 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.

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

d. A detailed evaluation against other prediction decision tree algorithms are performed, that provide a fair comparison to show the effectiveness of the proposed model.

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

The reminder of the paper is organized as follows: the next section reviewed the literature, which is followed by the description and working of the new model in section III. Section IV analyzes and discusses the result and finally section V concludes the paper with future directions and references.

2. RELEVANT WORK

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].

Irene Y.H Gu et al. [6], put forward a full automatic image analysis system for detection and analysis of snow/ice coverage on electric insulators of power lines using images which were captured by visual cameras in a remote outdoor laboratory test bed. Jinmei Pan et al. [7], put forth a passive microwave remote sensing techniques that detected wet snow in the south of china. Yajaira Mejia et al. [8], gave an approach for estimating the snowfall using neural networks on multi source remote sensing observations and ground based meteorological measurements. Melanie Wetzel et al. [9], projected a technique that supports the snowfall forecast and for the verification of radar limited mountainous terrain that includes matching the output parameters and graphics from high resolution mesoscale models to surface mesonets. Pascal Sirguey et al. [10], made use of ASTER and MODIS sensors, both on the TERRA platform by implementing the ARSIS concept so as to fuse the high spatial content of the two 250m spectral bands of MODIS into five 500m bands using wavelet based multi resolution analysis in the mountainous environment.

Michael A. Rotondi [11] illustrated a Markov chain models across eight national weather stations using historical data from the global historical climatology network to predict a ‘snow day’. Gail M. Skofronick Jackson et al. [12], in their research interpreted how instruments like the W-band radar of Cloudsat, Global

Precipitation Measurement Dual-Frequency Precipitation Radar ku- and Ka-bands, and the Microwave Imager can be used in the simulations of lake effect and synoptic snow events in order to determine the minimum amount of snow. Gail M. Skofronick Jackson et al. [13], demonstrated thresholds for detecting falling snow from satellite-borne active and passive sensors.

Andrea Spisni et al. [14], presented an operational chain developed in the Emilia-Romagna region to monitor snow cover and snow water equivalent over the area managed by the Regional Catchment Technical Service. Alberto Martinez Vazquez et al. [15] presented an algorithm using GB-SAR imagery for the automatic recognition and classification of snow avalanches. Jeremie Bossu et al. [16], made use of a structure, based on computer vision which detects the presence of snow or rain. Noel Dacruz Evora et al. [17], used brightness temperature data, provided by seven channels SSM/I aboard the Defense Meteorological Satellite Program F-11 and F-13 spacecrafts. Using which a modelling framework was put forth by combining passive microwave data, neural network based models and geostatistics for snow water equivalent retrieval and mapping. Hossein Zeinivand et al. [18], enforced a spatially distributed physically based model to detect snow and melting in the Latyan dam watershed in Iran.

Xiaolan Xu et al. [19], developed a model that can be used for both active and passive microwave remote sensing of snow. B.B Fitzharris et al. [20], presented three case studies on the usage of satellite imagery for mapping seasonal snow cover in New Zealand, and also explored the effectiveness of using AVHRR imagery in order to obtain the presence of snow, snow covered area and snow line elevation on the mountain ranges of New Zealand. Ashok N.Srivastava et al. [21], in their research discussed the results based on kernel methods for unsupervised discovery of snow, ice, clouds and other geophysical processes based on data from the MODIS instrument.

G. Singh et al. [22], developed a Radar Snow Index model to identify snow using SAR polarimetry techniques. In their research, full polarimetric L-band ALSOS-PALSAR data of snow cover area in Himalayan region have been analyzed based on various component scattering mechanism models and all model results are compared. Fan Ke et al. [23], developed a model to identify winter time heavy snow over Northeast China by using a inter annual increment prediction approach. Folorunsho Olaiya [24] investigated the use of artificial neural networks and decision tree algorithms in forecasting maximum temperature, rainfall, evaporation and wind speed using meteorological data collected from the city of Ibadan, Nigeria through Nigerian Meteorological Agency, Oyo state office. Manjeet Singh et al. [25] forwarded an attempt to develop an automatic technique for avalanche area identification and also its severity index. For detailed relevant work refer our earlier papers [38] [39] [41] [45-47].

3. ISLIQ-OC DECISION TREE ALGORITHM

The experimental implementation methodology of ISLIQ-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 gini indices for all the attributes; and 4) decision tree construction.

3.1. *I*SLIQ-OC Decision tree algorithm

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

$$G\_info = 1 - \sum_{i=1}^{M} P_i^2 \tag{1}$$

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 $G\_info_D)$ for each and every attribute "v_q"

$$G\_in\ fo_D = \sum_{j=1}^{N} P_j \left[1 - \sum_{i=1}^{M} P_i\right]^2 \tag{2}$$

7. Compute *GiniIndex* for each and every attribute "v_q"

$$Gini\ Index = G\_in\ fo - G\_in\ fo_D \tag{3}$$

8. The maximum *GiniIndex* is considered to be the best split point and becomes the root node.

$$Best\ Split\ point = Maximum(Gini\ Index) \tag{4}$$

9. Repeat Steps 6 through 8, generating leaf nodes in place of the root node until all leaf nodes contain the same class labels.

K-means clustering causes problems when the "k" parameter in k-means is set to a value that is considerably less than the inherent number of natural groupings within the training data. 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.

3.2. *Optimal* Decision Cluster

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.

4. RESULTS AND DISCUSSION

The data used for the present research is collected from www.wundergoruond.com [36]. The attributes considered here are Humidity, Temperature, Pressure, Wind Speed and Dew Point followed by Class Labels Snow and No-Snow. The proposed model has been tested on 20 international locations historical datasets. [35].

Table 1 summarizes the characteristics of the datasets, arranged in alphabetical order, presenting the number of instances, training instances, testing instances, attributes and classes. The comparison in terms of number of split points is presented in Table 2. Apparently, almost all results for ISLIQ-OC are better than those of SLIQ and ISLIQ. The results clearly show that the proposed ISLIQ-OC reduces splits when compared with other decision tree algorithms.

The comparison in terms of classification accuracy is presented in Table 3. The proposed method yielded an average accuracy of 89.63%, better, when compared with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS. For some of the cities, the accuracy levels are more for other algorithms when compared with ISLIQ-OC. But, on an average the ISLIQ-OC model outperforms when compared with other algorithms.

The comparison in terms of specificity is presented in Table 4. For some of the cities, the specificity levels are more for other algorithms when compared with ISLIQ-OC. But, on an average the ISLIQ-OC model outperforms when compared with other algorithms. The proposed method yielded an average specificity of 93.41%, better, when compared with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS.

The comparison in terms of error rate is presented in Table 5. Apparently, almost all error rate results for ISLIQ-OC are better than those of SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS. The proposed method yielded an average error rate of 10.36%, better, when compared with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS.

The comparison in terms of accuracy of ISLIQ-OC along with other non decision tree algorithms is presented in Table 6. Apparently, almost all error rate results for ISLIQ-OC are better than those of existing decision tree algorithms.

The comparison in terms of accuracy is presented in Table 7. Apparently, almost all accuracy results for ISLIQ-OC are better than those of those of existing decision tree algorithms.

Table 1
Dataset 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

Table 2
Split Points Comparison

City Name	*SLIQ*	*ISLIQ*	*ISLIQ-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
Accuracy Comparison Of Isliq-oc With Decision Tree Algorithms

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

Table 4
Specificity Comparison Of Isliq-oc With Decision Tree Algorithms

City Name	*SLIQ*	*SPM*	*SLGAS*	*ISLIQ*	*ISPM*	*ISLGAS*	*ISLIQ-OC]*
Aberdeen	93.24	91.2	92.19	93.69	93.24	66.99	90.47
Bangkok	96.36	94.8	95.45	98.53	98.74	98.32	98.95
Barcelona	97.59	97.8	97.87	98.83	98.83	98.28	98
Benton	71.82	73	75.08	73.07	72.28	75.97	69.45
Botswana	94.68	97.2	94.41	96.56	97.3	97.64	99.73
Brazil	79.52	77.1	80.7	77.53	79	77.75	74.05
Cairo	95.53	97	96.74	96.43	96.1	96.52	99.29
Chennai	82.36	81	82.36	75.27	78.74	76.38	90.47
Delhi	97.83	96.3	98.44	94.58	97.83	98.3	97.29
Eglinton	97.52	97.9	97.66	98.44	98.72	98.86	98.02
Humberside	95.49	98.4	98.36	97.13	97.13	97.13	97.54
Hyderabad	97.77	99.2	95.97	98.16	99.16	99.02	98.95
Iceland	96.51	93.3	95.48	93.67	96	94.32	94.96
Lahore	86.93	89.5	90.59	89.7	89.5	89.4	87.72
Manchester	95.69	95.9	92.94	96.9	94.01	95.83	95.02
Norway	95.92	98.1	97.78	97.92	97.71	97.42	94.78
Perth	96.73	96.9	98.8	96.87	96.73	96.53	96.87
Sellaness	78.74	83	85.45	93.02	92.42	88.46	90.87
Tiruptahi	98.06	97.9	98.06	97.93	98.06	97.93	99.53
Valley	93.86	95.3	93.72	93.91	94	91.45	96.27
Average	**91.1**	**91.6**	**92.05**	**92.2**	**92.3**	**92.87**	**93.41**

Table 5
Error Rate Comparison Of Isliq-oc With Decision Tree Algorithms

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

Table 6
Accuracy Comparison Of Isliq-oc With Non-decision Tree Algorithms

City Name	*Bayes Net*	*Naïve Bayes*	*Multilayer Perceptron*	*SMO*	*Simple Logistic*	*ISLIQ-OC*
Aberdeen	82.53	80.66	85.86	80.82	81.47	85.84
Bangkok	98.95	97.91	98.95	98.95	98.95	98.53
Barcelona	96.27	96.93	98.6	98.33	98.4	95
Benton	65.55	63.83	68.35	65.36	65.41	69.93
Botswana	99.27	99.4	99.4	99.2	99.13	98.21
Brazil	74.41	74.41	76.86	74.03	74.41	71.98
Cairo	95.83	96.67	97.65	97.52	97.65	91.6
Chennai	87.73	82.09	85.95	85.95	85.95	82.82
Delhi	86.16	85.29	90.48	89.28	88.95	95.74
Eglinton	94.49	95.69	95.75	95.18	95.69	89.24
Humberside	84.32	83.14	84.18	85.33	84.71	94.2
Hyderabad	99.31	99.31	99.31	99.31	99.31	97.6
Iceland	85.64	82	86.33	83.48	82.34	87.47
Lahore	83.85	80.09	86.96	87.21	86.97	84.65
Manchester	86.04	87.05	89.14	87.24	87.75	92.11
Norway	86.15	85.14	84.33	86.79	84.31	89.18
Perth	88.18	87.15	88.19	87.44	87.39	94.37
Sellaness	88.76	86.17	86.5	87.58	88.17	83
Tiruptahi	88.53	84.36	95.69	95.56	95.56	98.87
Valley	88.18	87.17	89.53	88.75	89.14	92.36

Table 7
Accuracy Comparison Of Isliq-oc With Existing Decision Tree Algorithms

City Name	*Decision Stump*	*J48*	*LMT*	*Random Forest*	*REP Tree*	*ISLIQ-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	98.53
Barcelona	98.33	98.33	98.4	98.33	98.33	95
Benton	63.29	68.48	68.52	66.18	68.24	69.93
Botswana	99.2	99.27	99.13	99.4	99.2	98.21
Brazil	68.44	65.85	67.24	67.99	65.92	71.98
Cairo	97.52	97.85	97.65	97.39	97.78	91.6
Chennai	85.95	85.95	85.95	85.35	86.02	82.82
Delhi	80.5	89.85	88.95	88.02	89.95	95.74
Eglinton	85.18	86.45	86.26	86.07	86.2	89.24
Humberside	91.35	91.28	91.65	92.32	91.55	94.2
Hyderabad	89.31	89.31	89.31	89.31	89.31	97.6
Iceland	81.77	87.35	87.47	85.53	86.44	87.47
Lahore	84.84	87.05	86.96	86.47	87.05	84.65
Manchester	84.72	88.63	88.06	87.24	88	92.11
Norway	88.41	86.49	86.11	86.31	86.69	89.18
Perth	91.42	93.21	90.73	92.69	91.76	94.37
Sellaness	78.83	80.24	81.27	80.09	79.87	83
Tiruptahi	95.56	95.56	95.56	95.42	95.56	98.87
Valley	88.75	89.14	89.73	89.01	89.07	92.36

5. CONCLUSION

Experimental results show that the ISLIQ-OC algorithm scales up well to both large and small datasets with large number of attributes and class labels. We compare our proposed method with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS decision tree algorithms in terms of the overall classification performance defined over four different performance measures namely split points, accuracy, specificity and error rate. Results on the snow/no-snow 20 international locations datasets show that:

a. the ISLIQ-OC decision tree reduces split points to a greater extent when compared with SLIQ and ISLIQ. The major contribution of this research is to reduce the split points, which is achieved.

b. the ISLIQ-OC decision tree outperforms in terms of classification accuracy over 20 international locations of snow/no-snow datasets. The proposed method yielded an average accuracy of 89.63%, better, when compared with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS.

c. the ISLIQ-OC decision tree outperforms in terms of classification specificity over 20 international locations of snow/no-snow datasets. The proposed method yielded an average specificity of 93.41%, better, when compared with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS.

d. the ISLIQ-OC decision tree outperforms in terms of classification error rate over 20 international locations of snow/no-snow datasets. The proposed method yielded an average error rate of 10.36%, better, when compared with SLIQ, SPM, SLGAS, ISLIQ, ISPM and ISLGAS.

e. the ISLIQ-OC decision tree outperforms in terms of classification accuracy over 20 international locations of snow/no-snow datasets when compared with decision and non-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)

[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)

[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)

[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

[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

[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. IRECOS, 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).