

Evaluating the Yield of Hybrid Napier Grass with Data Mining Techniques

Nadiammai.G.V
Ph.D (CS) Research Scholar
Karpagam University
Coimbatore, India – 641 021

Krishnaveni.S
Ph.D (CS) Research Scholar
Karpagam University
Coimbatore, India – 641 021

M.Hemalatha
Head, Dept. of Software
Systems & Research,
Karpagam University
Coimbatore, India – 641 021

ABSTRACT

Data Mining is the process of identifying the hidden patterns from large amount of data. It is commonly used in a marketing, surveillance, fraud detection and scientific discovery. In data mining, machine learning techniques are mainly focused as research through which we learnt to recognize complex and make intelligent decisions based on data. This paper involves the information about the yield of the hybrid grass from NBH1 to NBH11. The hybrid grass enhances the milk production in the states of Tamilnadu, Kerala, Karnataka, Andhra Pradesh, Orissa, and Maharashtra & Gujarat. It is well adapted to the soil and climatic conditions of Tamilnadu. In this paper, some of classification models are used to predict the yield of hybrid grass. They are NaiveBayes, J48, Rule Induction, Single Rule Induction, Decision Stump, ID3 and Random Forest.

Keywords

Data mining, machine learning, Naïve Bayes classifiers, J48, Rule Induction, Single Rule Induction, Decision Stump, ID3 and Random Forest

1. INTRODUCTION

Data mining is the extraction of hidden predictive information from large databases. Also it is a powerful technology with great potential to help companies focus on the most important information in their data warehouses [5]. Data mining tools predict future trends and behaviors, allowing businesses to make proactive, knowledge-driven decisions. The evolution of data mining began when business data was first stored on computers, continued with improvements in data access, and more recently, generated technologies that allow users to navigate through their data in the real time.

The Hybrid grass (*Pennisetum glaucum* × *Pennisetum Purpureum* Schumacher) is the major forage crop in the tropics and subtropics of the world. It has the potential to produce more dry matter per unit time than most other grasses (Hanna et al., 2004) [8]. It grows in areas of rainfall exceeding 100 mm. The hybrid Grass in a bunch type grass produces robust creeping Rhizome plants that have perennial growth habit in the tropics and subtropics and forms bamboo like clumps that grow up to 7 m in height. The hybrid grows best in regions with hot temperatures ranged from 25 to 40°C and growth stops when temperatures are below 10°C. It adopted well in altitude ranged from sea level to 2000 m and latitude

between 10°N and 20°S. Frost kills leaves and above ground stems, but the underground parts resumes growth at the beginning of the spring if the soil does not freeze.

In Tamilnadu, 65 lakh liters of milk are produced every day. Every year 4 percent growth in milk production is witnessed. During the eleventh plan period, 7-8 percent growth in milk production is desired (THE HINDU - 2007). The average milk yield in India is 987 kg/lactation, whereas it is 9291 kg in Israel, 7038 kg in U.S.A and 6273 kg in Denmark. In order to enhance the milk production and meet the demand, cross breed cows should be fed with nutritious green fodder throughout the year (Hedge 2007) [9]. The annual demand of green fodder to feed the animal population of 24,621,161 is 83.75 million tons. But at present the supply of green fodder is 12.68 million tons only, leaving a deficit of 71% (DAHUS 2004) [4]. Presently the Pearl Napier Hybrid grass is grown throughout India, bearing some hilly areas where frost is a common phenomenon.

2. RELATED WORKS

The hybrid grass yields surpass those of Rhodes Grass (*Chloris Gayana*), *Setaria* (*Setaria Sphacelata*) and Kikuyu Grass (*Pennisetum Clandestine*) which are popular pasture grasses but which yield between 5 to 15 tones of dry matter per year. High dry matter yields up to 85 tones/ha has been recorded elsewhere in the tropics (Orodho, 2006) [15]. Dry matter yield alone, however, is of limited value if it is not closely related to the dry Matter intake of the animals. As the grasses tolerates frequent defoliation, under good weather condition, it can be cut every 6-8 weeks giving up to 8 cuts in a year, depending on fertilizer applications, rainfall amount and distribution.

After Burton (1944) the interspecific hybrids first produced in the United States, then where produced in several other countries like India (Krishnaswamy and Raman, 1949) [13], South Africa (Gildenhuys, 1950) [7], Pakistan (Khan and Rahman, 1963), Australia (Pritchard, 1971) [17] and Nigeria (Aken'Ova and Chedda 1973) [1]. The main goal of crossing these species was to produce high yielding, high quality perennial fodder hybrids combining pearl millet's forage quality, non shattering nature, and ability of establish readily, as well as the perennial aggressive nature of Napier grass.

The interspecific hybrid is a triploid (AA'B genomes) with $2n=3x=21$ chromosomes (7 chromosomes from Pearl Millet and 14 chromosomes from Napier Grass). Triploid hybrids resulting from the interspecific cross are usually

highly variable because of the heterozygosis of Napier grass, even if the pearl millet parent is an inbred (Hanna et al. 2004) [8]. The largest collection of Napier grass germplasm is present at the nursery in Tifton, GA where over 100 plant introductions from around the world and breeding line material from intra and interspecific crosses are maintained.

The hybrids generally show high heterosis for fodder yield and quality, and thus are high yielding and more acceptable than the Napier grass parent (Osgood et al., 1997) [16]. However, these hybrids are sterile and need to be propagated vegetative, which puts a major limitation on their easy distribution to farmers.

Babu et al. (2009) [2], Randomly Amplified Polymorphic DNA (RAPD) and Inter-Simple Sequence Repeat (ISSR) markers were used to detect the DNA polymorphism among thirty Napier grass collections of wide geographical distribution maintained at TNAU, Coimbatore.

In this paper we investigate application of Naïve Bayes, J48, Rule Induction, Single Rule Induction, Decision Stump, ID3 and Random Forest classifier and compares these algorithms performance based on Hybrid Napier Grass. According to the variable definition for the TNAU-Forage Crops records of dataset, this Hybrid Napier Grass has the varieties such as NBH1 to NBH11, CN92, CO3 and TNCN014. It includes the labels of plant height, leaf length, leaf breadth, and leaf width, number of leaves, leaf stem ratio and number of tillers. In this paper, we had taken only three labels. They are plant height, number of leaves and number of tillers for varieties from NBH1 to NBH11.

3. CLASSIFICATION TECHNIQUES

A major focus of machine learning [6] research is automatically learn to recognize a complex patterns and make intelligent decisions based on data. Hence, machine learning is closely related to fields such as statistics, probability theory data mining, pattern recognition, artificial intelligence, adaptive control, and theoretical computer science.

3.1 Naive Bayesian Classifier

A Naïve Bayesian classifier [10] is a simple probabilistic classifier based on applying Bayesian theorem (from Bayesian statistics) with strong (naïve) independence assumptions. Using Bayesian theorem, we write

$$p(C | F_1, \dots, F_n) = \frac{p(C) p(F_1, \dots, F_n | C)}{p(F_1, \dots, F_n)}$$

Advantages

- It is fast, highly scalable model building and scoring
- Scale linearly with the number of predictors and rows
- Build process for Naïve Bayes is parallelized

- Induced classifiers are easy to interpret and robust to irrelevant attributes

- Uses evidence from many attributes, the naïve Bayes can use for both binary and multi-class classification problems

3.2 J48 Decision Tree Classifier

J48 is a simple tree [18], it creates a binary tree. C4.5 builds decision trees from a set of training data which is like an ID3, using the concept of information entropy.

The algorithm

- Check for base cases
- For each attribute 'c' find the normalized information gain from splitting on 'c'
- Let c_{best} be the attribute with the highest normalized information gain
- Create a decision node that splits on a c_{best}
- Recurse on the sub lists obtained by splitting on c_{best} , and add those nodes as children of node

Advantages

- Gains a balance of flexibility and accuracy
- Limits the number of possible decision points
- It provides a higher accuracy

3.3 Rule Induction

The Rule Induction methods [11] generate only 'correct' rules, measured by the accuracy formula p/t. Rule induction is an area of machine learning in which formal rules are extracted from a set of observations. The rules extracted may represent a full scientific model of the data, or merely represent local patterns in the data. This operator works similar to the propositional rule learner named Repeated Incremental Pruning to Produce Error Reduction (RIPPER, Cohen 1995). Starting with the less prevalent classes, the algorithm iteratively grows and prunes rules until there are no positive examples left or the error rate is greater than 50%.

The algorithm

- It always starts with an empty rule which covers all examples, and then restricts it by adding new conditions (attribute-value pairs)
- Until it covers only examples of the desired target class
- At each stage of adding the condition to the rule, the best attribute-value pair in terms of p/t (accuracy) is chosen

- If there are more attribute-value pairs with the same value of p/t , then the one with greatest coverage is chosen

3.4 Single Rule Induction (Single Attribute)

This operator concentrates on one single attribute and determines the best splitting terms for minimizing the training error. The result will be a single rule containing all these terms.

3.5 Decision Stump

A decision stump is a machine learning model with a one-level decision tree [19] also called 1-rules. It includes one root node that is connected to the terminal nodes. A decision stump makes a prediction based on the value of just a single input feature. That is a small decision tree with only one single split. This decision stump works on both numerical and nominal attributes. In nominal features, we can have a stump with two leaves that corresponds to some particular category, and remaining leaf trends to all the other categories. In case of binary features two schemas are identical and missing value is taken as another category [14]. In continuous features, some threshold value is selected, and the stump contains two leaves for values below and above the threshold value. However, rarely, multiple thresholds may be chosen and the stump therefore contains three or more leaves.

3.6 ID3

It is used to generate a decision tree [12]. It is also a precursor to the C4.5 algorithm. It is a mathematical algorithm for building the decision tree, introduced by J. Ross Quinlan (1979).

The Algorithm

- Builds the tree from the top down, with no backtracking
- Information Gain is used to select the most useful attribute for classification

Advantages

- Understandable prediction rules are created from the training data.
- It builds the fastest and short tree.
- Only need to test enough attributes until all data is classified.
- Finding leaf nodes enables test data to be pruned, reducing number of tests.
- Whole dataset is searched to create tree.

3.7 Random Forest

Random forest [3] is a collection classifier, consists of many decision trees and outputs the class.

The Algorithm

- Let the number of training cases be N , and the number of variables in the classifier be M .
- The number m of input variables are used to determine the decision at a node of the tree; m should be much less than M .
- Choose a training set for this tree by choosing N times with replacement from all N available training cases. Use the rest of the cases to estimate the error of the tree, by predicting their classes.
- For each node of the tree, randomly choose m variables on which to base the decision at that node. Calculate the best split based on these m variables in the training set.
- Each tree is fully grown and not pruned

Advantages

- It produces a highly accurate classifier and learning is fast
- It runs efficiently on large data bases
- It can handle thousands of input variables without variable deletion
- It computes proximities between pairs of cases that can be used in clustering, locating outliers or (by scaling) give interesting views of the data
- It offers an experimental method for detecting variable interactions

4. DATASET COLLECTION

This work used data which is produced by the Tamil Nadu Agriculture University (Forage Crops Department), Coimbatore. The dataset contains the details about the Hybrid Napier Grass records of 2006. It has the varieties such as NBH1, NBH2, NBH3, NBH4, NBH5, NBH6, NBH7, NBH8, NBH9, NBH10 and NBH11, CN92, CO3 and TNCN014. It includes the attributes of plant height, leaf length, leaf breadth, and leaf width, number of leaves, leaf stem ratio and number of tillers. The total records are 3430. In this we are using only 312 cases for our study. They are depends upon the labels such as plant height (104), number of leaves (104) and number of tillers (104) for varieties from NBH1 to NBH 11. The table (1) shows the details about the varieties.

Table 1. Varieties details in dataset

Varieties	Origin	Parentage
NBH 1	Vellayani	TNSC – 4 × FD 471
NBH 2	Vellayani	HES – 4 × FD 467
NBH 3	Coimbatore	Co 8 × FD 440

NBH 4	Coimbatore	Co 8 × FD 467 – 11
NBH 5	Coimbatore	Co 8 × FD 461 – 6
NBH 6	Coimbatore	Co 8 × FD 467 – 6
NBH 7	Coimbatore	Co 3 National check
NBH 8	IARI	NB - 21 National check
NBH 9	Rahuri	RBN 2008
NBH 10	Dharwad	DHN – 20
NBH 11	Dharwad	DHN – 6

5. PERFORMANCE ANALYSIS

This paper deals with the performance of seven classification algorithms namely Naive Bayesian, J48 Decision Tree Classifier, Rule Induction, Single Rule Induction, Decision Stump, ID3 and Random Forest. Data has been collected from Tamil Nadu Agriculture University (Forage Crops Department), Coimbatore. The dataset is recorded into three different scenarios

- Based On Leaves Information
- Based On Height Information
- Based on Tillers Information

The dataset contains the details about the Hybrid Napier Grass records. It has the varieties such as NBH1, NBH2, NBH3, NBH4, NBH5, NBH6, NBH7, NBH8, NBH9, NBH10 and NBH11, CN92, CO3 and TNCN014, and the labels are plant height, leaf length, leaf breath, and leaf width, number of leaves, leaf stem ratio and number of tillers. In this we are using only three labels such as plant height, number of leaves and number of tillers for the varieties NBH1 to NBH 11. In the total record set we are used 104 instances for Leaves Information, 104 instances for Height Information and 104 instances for Tillers Information.

5.1 Based on Naive Bayesian Classification algorithm

The dataset includes the information about the growth of the Hybrid Napier Grass for the varieties NBH1 to NBH11. Every variety includes 104 instances for all attributes such as leaves, height and tillers. Naive Bayes algorithm is used for specifying all the correctly and incorrectly classified instances along with number of records and accuracy percentage for the same. The correctly classified Naive Bayes Classifiers' percentages of the attributes are listed in the table (2).

Figure 1 shows the growth of Hybrid Napier Grass Dataset. In this the NBH 1 crop variety takes 82.69 %, 80.12 % and 83.54 % for leave, height and tiller attributes respectively. The NBH 2 crop variety takes 79.81 %, 79.55 % and 89.42 % for leave, height and tiller attributes

respectively. Likewise the correctly classified percentages of all the varieties are shown in this figure. Finally, this figure shows that the NBH6 variety is the best in leaves and tillers. NBH7 is the best in height.

Table 2. Dataset Classification based on Naive Bayes Algorithm

Crop Varieties	Leaves	Height	Tillers
NBH1	82.69	80.12	83.54
NBH2	79.81	79.55	89.42
NBH3	79.88	82.32	85.58
NBH4	78.69	81.90	90.27
NBH5	84.62	79.96	88.46
NBH6	86.79	85.33	98.08
NBH7	78.96	88.85	97.12
NBH8	81.96	82.23	96.15
NBH9	75.00	79.96	97.31
NBH10	78.85	83.44	91.58
NBH11	79.81	82.66	94.31

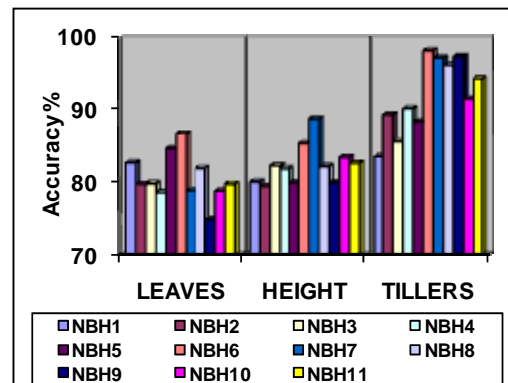


Figure 1. Based on Naive Bayesian Classification Algorithm

5.2 Based on J48 Decision Tree Algorithm

The dataset includes the information about the growth of the Hybrid Napier Grass for the varieties NBH1 to NBH11. Every variety includes 104 instances for all attributes such as leaves, height and tillers. J48 algorithm is used for specifying all the correctly and incorrectly classified records along with accuracy percentage for the same. The correctly classified percentages of the attributes are listed in the table (3).

Table 3. Dataset Classification based on J48 Algorithm

Crop Varieties	Leaves	Height	Tillers
NBH1	98.12	98.08	97.12
NBH2	98.08	97.12	97.27
NBH3	99.04	97.88	94.23
NBH4	99.04	96.76	95.19
NBH5	96.15	97.92	96.15
NBH6	99.24	92.24	99.12
NBH7	95.19	99.08	97.12
NBH8	93.31	95.48	96.15
NBH9	97.12	99.04	97.31
NBH10	95.19	95.19	98.08
NBH11	96.23	96.15	94.31

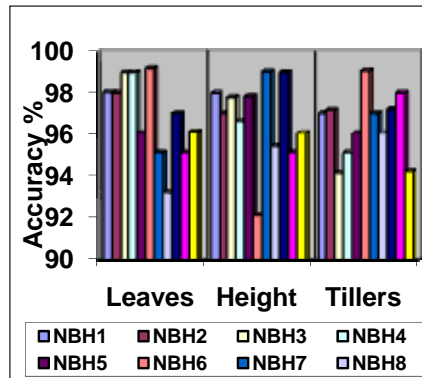
**Figure 2. Based on J48 Classification Algorithm**

Figure 2 shows the growth of Hybrid Napier Grass Dataset based on J48 algorithm. In this the NBH 1 crop variety takes 98.12 %, 98.08 % and 97.12 % for leave, height and tiller attributes respectively. The NBH 2 crop variety takes 98.08 %, 97.12 % and 97.27 % for leave, height and tiller attributes respectively. Likewise the correctly classified percentages of all the varieties are shown in this figure. Finally, this figure shows that the NBH6 variety is the best in leaves and tillers. NBH7 is the best in height.

5.3 Based on Rule Induction and Single Rule Induction Algorithms

The dataset includes the information about the growth of the Hybrid Napier Grass for the varieties NBH1 to NBH11. If NBH1 is used as a label for the attribute leaves, the classes will be specified as [NBH1, 12, 11, 10, 9 and 8]. Here 12, 11, 10, 9 and 8 are the number of leaves.

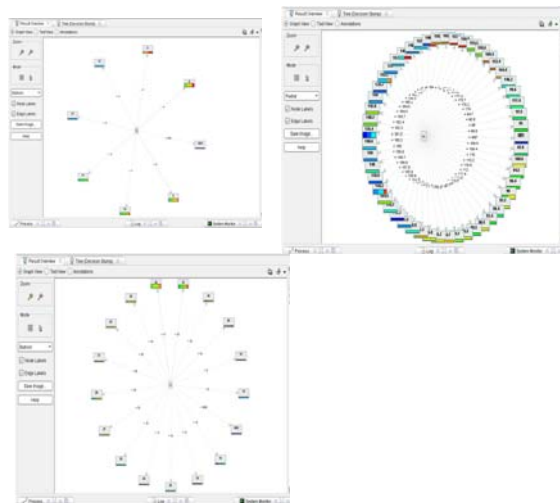
Every variety includes 104 instances for all attributes such as leaves, height and tillers. The correctly trained instances are listed in the table (8) shows the values for Rule Induction and Single Rule Induction Algorithms. In case of rule induction the correctly trained instances are 87, 15 and 89 for the attributes leaves, height and tillers respectively. In case of single rule induction the correctly trained instances are 62, 100 and 63 for the attributes leaves, height and tillers respectively.

Table 4. Dataset Classification based on Rule Induction and Single Rule Induction Algorithm

Attributes	Rule Induction	Single Rule Induction
Leaves	87	62
Height	15	100
Tillers	89	63

5.4 Based on Decision Stump Algorithm

The dataset includes the information about the Hybrid Napier Grass for the varieties NBH1 to NBH11. Every variety includes 104 instances and NBH1 is used as a label for all attributes such as leaves, height and tillers.

**Figure 3. Based on Decision Stump for leaves, height and tillers**

The figure 3 shows the overall view of the classes. For leaves attribute, NBH6 acts as a root node with the combination of values 8, 9, 10, 11, 12, 13 and 14. It has the classes namely 9, 9, 9, 10, 11, 12 and 12. Here we have four classes only, because the values split based upon these classes and the majority goes to 12 and 9. The root node takes NBH1 as one among the class. NBH6's total instances 104 had split as 9, 31, 27, 16, 13, 4 and 4 as per size. For height attribute, NBH 7 acts as a root node. In this we can see the classes with color change for the different type of classification. Each value has some

specific color. According to the tillers attribute, NBH6 is taken as root node and classification are made based upon the label NBH1. It has the combination of values 8, 9, 10, 15, 16, 17, 18, 24, 25, 28, 29, 30, 31, 32, 33 and 34 with classes 10, 14, 11, 15, 24, 25, 27, 28, 29, 30, 31, 32, 33 and 34. It does not include any sub trees but it provides a clear view.

5.5 Based on ID3 Algorithm

The dataset includes the information about the Hybrid Napier Grass for the varieties NBH1 to NBH11. Every variety includes 104 instances and NBH1 is used as a label for all attributes such as leaves, height and tillers.

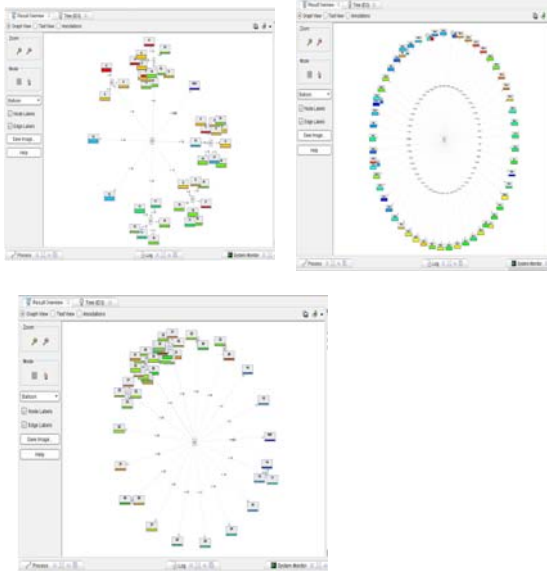


Figure 4. Based on ID3 Algorithm for leaves, height and tillers

In the figure 4 the ID3 clearly shows all the classes, subclasses, values and its sizes. For leaves attribute, NBH6 is taken as the root node with 10 values, 9 classes and size 27 with 8=1, 9=13, 10=10, and 12=3. It includes another sub tree NBH7 with values 7, 8, 9, 10, 11, 12, 13, 14, etc. Likewise it classifies the remaining values with NBH1 as a label. For height attribute, NBH1 is placed as label and NBH7 is taken as root node. Totally it classifies 105 samples. From this 53 values are shown in the above graph. NBH7 includes a sub tree NBH3 with size 6 (i.e. six values). Among this, the values has sub tree named NBH2. Likewise, it classifies the remaining values based on the label NBH1. For tillers attribute, NBH6 is taken as root node and classifications are made based upon the label NBH1. It has the combination of values 8, 9, 10, 15, 16, 17, 18, 24, 25, 28, 29, 30, 31, 32, 33 and 34, classes 10, 14, 11, 15, 24, 25, 27, 28, 29, 30, 31, 32, 33 and 34. The values 10, 25, 30, 31 has the sub tree namely NBH2[s=3], NBH2[s=4], NBH11[s=6] and NBH4[s=37]. ID3 gives the detailed method of classification.

5.6 Based on Random Forest Algorithm

The dataset includes the information about the Hybrid Napier Grass for the varieties NBH1 to NBH11. Every

variety includes 104 instances and NBH1 is used as a label for all attributes such as leaves, height and tillers. Random Forest includes model 1 to 10. Each model had different type of classification based on the values by keeping NBH1 as label.

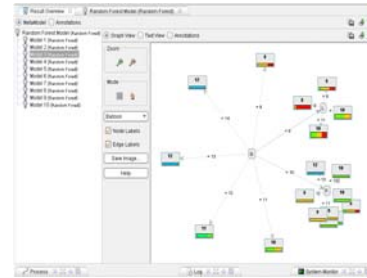


Figure 5. Based on Random Forest Algorithm for leaves

The figure 5 shows that the third model of Random Forest algorithm for the attribute leaves, in that NBH1 acts as label with the combination of values 12, 11, 10, 9 and 8. Based on these values, classifications are made. In model 3, 7, 9 and 10, NBH6 acts as root node and most of them are classified based on the values 9 and 10. Value 9 has the sub tree NBH11 with values 8, 9, 10 and 11, classes 8, 9, 10 and 10, size are 9[9=1, 8=8], 6[10=1, 9=3, 8=2], 12[10=7, 9=5] and 17[10=7, 9=5, 8=5]. The value 10 has the sub tree NBH10 with size 27 and the values are 10, 11, 12, 13, and 13. In turn 10 has the sub tree NBH3 with values 10 and 9 and classes 9=4 and 10 =4. Likewise all the other values are classified.

For height attribute, it had classified based on the values 150, 150.2, 146.5, 90.5, 148.6, 94.7 and 95. Models 1, 6, 9 and 10 are classified based on the value 150 and it had different sizes. For model 1[150=8], model 6[150=10], model 9[150=7] and model 10[150=8]. Likewise all the other values are classified. For the attribute tillers had classified based on the value 32 for all the models (1 to 10). It has different sizes for model 1[32=41], model 2[32=43], model 3[32=32], model 4[32=31], model 5[32=38], model 6[32=27], model 7[32=36], model 8[32=34], model 9[32=33] and model 10[32=38]. For example model 1: 32 {NBH1=0, 14=1, 10=1, 15=1, 11=1, 24=2, 25=0, 30=13, 33=16, 32=41, 27=1, 28=4, 31=22, 29=0, 34=2}.

6. CONCLUSION

This paper deals with performance of seven classification algorithms namely Naive Bayesian Classifier, J48 Decision Tree Classifier, Rule Induction, Single Rule Induction, Decision Stump, ID3 and Random Forest. The dataset contains the details about the Hybrid Napier Grass records. It has totally 14 varieties and 7 attributes. In that we took 11 varieties and three attributes for our analysis. The attributes are plant height, number of leaves and number of tillers and the varieties are from NBH1 to NBH11. In this analysis, finally we had found that the NBH6's yield is best for the attributes leaves and tillers. The NBH7's yield is best for the attribute height.

When using Naive Bayes algorithm it shows the correctly classified percentage values for all varieties. J48 gives

more accuracy than Naïve Bayes classification algorithm. Single rule induction gives result based upon any one of the variety (attribute) and determines the best splitting terms for minimizing the training error. The rule induction generates only correct rules based on the accuracy. The Decision Stump gives the overall view (values and classes). The ID3 algorithm gives the detailed view (values, classes, sub trees and their details). Random Forest algorithm shows the different models and each model gives different results. From that the best one is taken into account. By comparing all these, Random Forest outperform than other algorithms.

7. ACKNOWLEDGEMENT

We thank the Karpagam University and our guide Dr.M.Hemalatha for the Motivation and Encouragement to make this research work as successful one.

8. REFERENCES

- [1] Aken'ova M.E. and H.R.Chheda. 1981. Morphology, cytology and forage potential of Pennisetum americanum (1) K.Schum.x P.Purpureum schum. Amphidiploids Euphytica 30, pp. 397-404. Publishers, Dortrecht, Netherlands. Pp. 343
- [2] Babu,C., Sundramoorthi,J., Vijayakumar.G and Ganesh Ram.S, 2009. Analysis of Genetic Diversity in Napier Grass (Pennisetum purpureum Schum) as detected by RAPD and ISSR Markers. J.Plant Biochemistry & Biotechnology Vol 18(2), 181-187.
- [3] Breiman, Leo (2001). "Random Forests". Machine Learning 45 (1): 532. doi:10.1023/ A:1010933404324
- [4] DAHUS-Director of Animal Husbandry and Veterinary Services Report, 2004, Chennai, Tamil Nadu.
- [5] Data Mining and Data Warehousing available at:<http://databases.about.com/od/datamining/g/gclassification.htm>
- [6] Domingo's, Pedro & Michael Pazzani (1997) "On the optimality of the simple Bayesian classifier under zero-one-loss". Machine Learning, 29:103-137.
- [7] Gildenhuis, P.J.1950. A new fodder grass, Farming in South Africa, 15, 189-191.
- [8] Hanna,W.W., C.J.Chaparro, B.W.Mathews, J.C.Burns, L.E.Sollenberger, and J.R.Carpenter. 2004. Perennial Pennisetums. Pp.503-535. In L.E.Moser, B.L.Burson, and L.E.Sollenberger (Ed.). Warm season (C4) grasses. ASA/CSSA/SSSA,Madison,WI.
- [9] Hedge, N.G., 2007. Transfer of technology for forage production, BAIF Development Research Foundation, Pune (Personal Communication)
- [10] Naïve Bayes classifier, available at http://en.wikipedia.org/wiki/Naive_Bayes_classifier
- [11] Rule induction, Available at http://en.wikipedia.org/wiki/Rule_induction
- [12] DMS Tutorial - Decision trees, Available at http://dms.irb.hr/tutorial/tut_trees.php-id3
- [13] Krishnaswamy.N. and V.S.Raman, 1949. A note on the Chromosome numbers of some economic plants of India.Curr.Sci., 18.376-375.
- [14] Loper, Edward L.; Bird, Steven; Klein, Ewan (2009). Natural language processing with Python. Sebastopol, CA: O'Reilly. ISBN 0-596-51649-5.
- [15] Orodho.A.B. 2006. The role and importance of Napier grass in the smallholder dairy industry in Kenya, P.O.box 1667, kitale-30200, Kenya.FAO edition.http://www.fao.org/ag/AGP/AGPC/doc/Newpub/napier/napier_kenya.html.
- [16] Osgood, R.V.Hanna.W.W, Chaparro and B.W.Mathews.1997.Hybrid seed Production of Pearl millet x Napier grass triploid hybrids. Crop Science 37(3):998-999. {a}USDA-ARS, Coastal Plain Exp.Stn, Tifton, GA, USA.
- [17] Pritchard, A.J (1971).The hybrid between Pennisetum typhoid's and P.Purpureum as a potential forage crop in South Eastern Queensland. Tropical Grasslands, Vol.5, No.1, PP.35-39.
- [18] Quinlan, J.R.C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.
- [19] Wayne Iba and Pat Langley, (1992). Induction of One-Level Decision Trees. Proceedings of the Ninth International Conference on Machine Learning.