Indian Musical Instrument Recognition using Modified LPC Features

Satish R.Sankaye
Department of CS & IT
Dr.Babasaheb Ambedkar
Marathwada University,
Aurangabad

Suresh C.Mehrotra
Department of CS & IT
Dr.Babasaheb Ambedkar
Marathwada University,
Aurangabad

U.S. Tandon
Department of Physics,
College of Natural and
Computational Sciences, Haramaya
University, Dire Dawa, Ethiopia

ABSTRACT
Indian Classical Music is considered very diverse and distinct area of music across the globe. It has its indistinct melodies especially made up of unique musical instruments. It uses a wide variety of Musical Instruments to achieve this feat. In last two decades, researchers are actively associated with human perception towards the study of Musical Instruments. In this paper, we have proposed an innovative method to classify the Indian Musical Instrument Recognition (IMIR) technique using the Modified Linear Predictor Coefficient (LPC) features. The Classification algorithm has adopted Linear Discriminant Analysis (LDA). The proposed method has been tested with nine kinds of musical instruments. The research project involved the identification of musical sounds with experimental results using the present technique which has an accuracy of 93.04%.

General Terms
Indian Musical Instrument Recognition (IMIR), Linear Predictor Coefficient (LPC), Linear Discriminant Analysis (LDA), Best First Decision Tree, Temporal Features, Spectral Features, Monophonic, Polyphonic audio signal.

Keywords
Indian Musical Instrument Recognition (IMIR), Linear Predictor Coefficient (LPC), Linear Discriminant Analysis (LDA), Best First Decision Tree.

1. INTRODUCTION
Indian classical music has longest unbroken tradition of Indian Musical Instruments. These prestigious traditions have endowed many musical wonders. Indian musical instruments have no alternatives in terms of their special melodious quality and credibility in the world of music. They have retained unique place with their sound quality, intricacy, mechanics, verisimilitude to the human voice that involves deep emotions and powerful passionate appeal through the melodies. The instrumental melodies can strike the heart of listeners with its melody and sentimental appeal. This unique feature of Indian musical instruments has attracted numerous researchers to study it widely. The present research project is an attempt to put forward a novel approach for the identification of the Indian Musical Instrument as Indian Musical Instrument Recognition (IMIR).

Globally, Musical Instrument recognition (MIR) context can be alienated into three classes: Monophonic Isolated Note Instrument Recognition, Monophonic Musical Phrases Instrument Recognition, and Polyphonic Musical Instrument Recognition [1]. Mostly MIR is based on physical structure of a musical instrument on its temporal and spectral levels. Its Spectral envelope features are derived from Linear Predictor

Coefficient (LPC), Mel Frequency Cepstral Coefficient (MFCC), Pitch Periodogram and Absolute Amplitude Envelope [2].

Researchers are keen to study and analyze the human perception of the Musical Instrument Identification. They are curious for Retrieval, Classification, Recognition and management of large sets of music that is known as Music Information Retrieval. Our research has its major focus on evolving an innovative method to recognize Nine Indian Musical Instruments with their distinct melodies.

The present article contains 6 Sections elaborating IMIR. In Section 2 we discuss reviews of latest research in the area of Musical Instrument Recognition (MIR). Section 3 deals with the theoretical concepts related to the research topic however Section 4 describes the proposed architecture of Indian Musical Instrument Recognition System and the detailed analysis of data set. Section 5 contains the results and its analysis. And the paper is concluded in Section 6 with its proposition and future direction.

2. RESEARCH REVIEW
Antti Eronen and Anssi Klapuri [3] proposed a system for Pitch Independent Musical Instrument Recognition. They studied about 30 musical instruments belonging to string, brass and woodwind families. The correct instrument family was recognized with 94% accuracy and individual instruments in 80% of cases. They have considered a wide set of acoustic, temporal and spectral features of sounds in instrument recognition.

Giannoulis Dimitrios and Anssi Klapuri [4] suggested method using local spectral features and Missing-feature technique for musical instrument recognition in polyphonic audio signals. They recommended a mask estimation technique based on the assumption that the spectral envelopes of musical sounds tend to be slowly-varying as a function of log-frequency. This technique outperformed the reference methods by a wide margin, indicating that the missing feature approach which provided a significant robustness improvement in processing polyphonic audio.

Shah, Jashmin K., et al. recommended a method to classify the given frame of a speech waveform as voiced speech or unvoiced speech [5]. They have presented two novel approaches using acoustical features and pattern recognition. With the first approach they have used Mel frequency cepstral coefficient with Gaussian mixture model classifier in which they have obtained approximately 90% identification accuracy. The other method is based on LPC coefficient and reduced dimensional LPC residual with Gaussian mixture

model classifier, which resulted in 92% identification accuracy.

Zlatintsi, Athanasia, and Petros Maragos [6] explored the nonlinear methods of fractal theory and proposed the use of a multiscale fractal feature for structure analysis of musical instrument tones. They have used musical instruments such as *Double Bass, Bassoon, Bb Clarinet, Cello, Flute, French, Horn* and *Tuba* for their study and analysis. They computed the short-time MFDs of the tones using 30 ms segments of the full duration. The results have shown that if the MFDs fused with the MFCCs, it improved the recognition accuracy. It accomplished a reduction of error up to 32%.

Hai, Jiang, and Er Meng Joo [7] presented new strategy for feature extraction of speech recognition. They used LPC + ΔLPC coefficients feature vectors and transform it into another kind of simple representation with LDA. They have implemented improved LPC feature method, for a high performance speech recognition system with Hidden Markov Model very successfully.

Gunasekaran and K. Revathy [8] have explored in depth a classifier combination approach for the Indian instrument classification task. They studied over a diverse classifier pool that included K-Nearest Neighbor, Gaussian Mixture Model and Multi-Layer Perceptron classifiers. Their test results of individual classifiers were in the range of around 86.7% to 90.6%. An improvement in the result was achieved with sum of the based classifier up to 93.60%.

3. THEORTICAL CONCEPT

3.1 Linear Predictive Coefficient

The general principle behind Linear Predictive Coefficient (LPC) [9] is an approach to a source signal, analyzed and resynthesized according to a source-filter model of human voice. Parametric representation of a spectrum using linear predictor Coefficient is a powerful technique in speech processing which is now prominently used in Musical Instrument Retrieval System as well.

LPC is computed from the autocorrelation vector using a Levinson or a Durbin recursion method. The signal at time *t* is approximated with linear combination of signal values in previous moments, which can be mathematically expressed as follows:

$$\widetilde{\mathbf{f}_t} = \sum_{k=1}^{p} \mathbf{a}_k \mathbf{f}_{t-k}$$

.

In the above equation **t** is discrete time moment, $\mathbf{f_t}$ is original signal, $\mathbf{a_k}$ is the Linear Predictive Coefficient; $1 \leq k < p$ and p is number of Linear Predictive Coefficients.

3.2 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is widely used in face recognition, mobile robotics, object recognition and musical Instrument Classification [10] [11] [12]. LDA is one of the common tools for multi-group data classification and dimensionality reduction. It is a technique with aim to transform and attempt to minimize the ratio of the within-class scatter to the between-class scatter that guarantee the maximum classification separability [13].

LDA searches for those vectors in underlying space that best discriminate among classes. It creates a linear combination which yields the largest mean difference between the desired classes. Mathematically speaking, for all the samples of all classes, we define two measures:

1) One is called within-class scatter matrix. The within-class scatter matrix SW [14] is computed by the following equation:

$$S_W = \sum_{i=1}^{c} S_i$$

where

$$S_i = \sum_{\mathbf{x} \in D_i}^{n} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T$$

(scatter matrix for every class)

and $\mathbf{m_i}$ is the mean vector

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{x} \in D_i}^{n} \mathbf{x}_k$$

2) The other is called between-class scatter matrix. The **between-class scatter** matrix S_B is computed by the following equation:

$$S_B = \sum_{i=1}^{c} N_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T$$

Where **m** is the overall mean, and $\mathbf{m_i}$ and $\mathbf{N_i}$ are the sample mean and sizes of the respective classes.

3.3 Best First Decision Tree

Best-first decision tree learning [15] is a kind of decision tree learning which has almost all properties of standard decision learning. A decision tree algorithm is especially good for classification learning if the training instances have errors (i.e. noisy data) and attributes have missing values. It builds a decision tree using a best-first search strategy [16].

Best-first decision tree is constructed in the divide and conquer fashion [17]. Each non-terminal node in a best-first decision tree tests an attribute which is assigned a classification. Three important things must be considered during the process of construction. The first one is to find the best attribute to split at each node. The second is to find the appropriate node in the list (i.e. all nodes are the candidates for splitting) that is to be expanded later. The third is to make the decision when to stop growing trees.

4. PROPOSED ARCHITECTURE

The present research paper proposes Indian Musical Instrument Recognition (IMIR) technique using the Modified Linear Predictor Coefficient Feature Vector with the help of Linear Discriminant Analysis Classification Method. The proposed architecture is as shown in Fig 1.

We have collected the sound samples of different Indian Musical Instruments. The samples are directly played by the instrumentalist; some of them are Pre-Recorded Compact disc of live performances, and some samples are collected from internet sources. The collected Sound Samples are recorded in the stereo with bit depth of 16 bits and sampling rate 44.1 kHz in .wav file format. Fig 2, shows the waveform of the sound samples of *Ghungaroo, Tabla* and *Shehnai* respectively. The Modified LPC Feature in the set of 15 values is computed as discussed below:
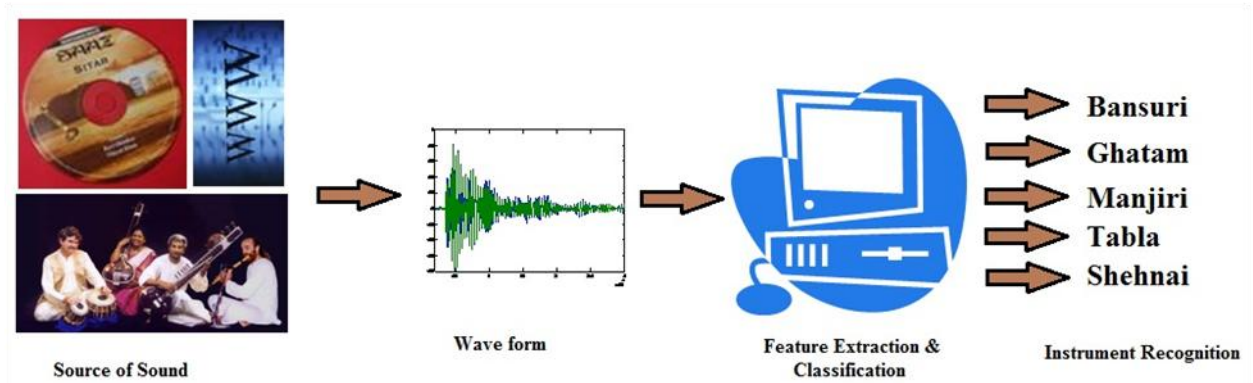
Fig 1: Diagrammatic representation of the proposed work to recognize the Indian Musical Instrument
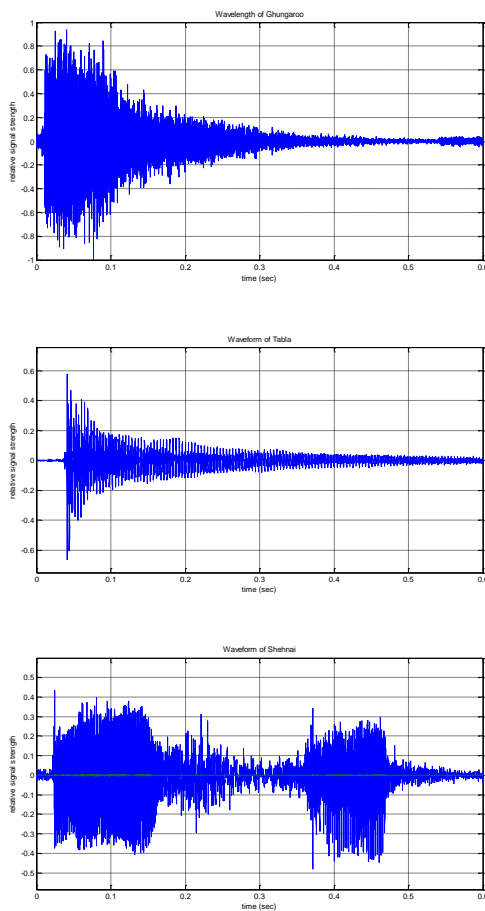

Fig 2: Waveform for the Musical Instrument *Ghungaroo, Tabla and Shehnai*

4.1 Modified LPC Feature Set

The modified LPC Feature set is computed as follow:

Step 1: The Waveform of sound samples is collected in the feature Vector X_i. The feature Vector X_i has been smoothened by computing an LPC feature set, A = [1 A(2) .... A(N+1)] of Nth order linear predictor using

$$n = [2+fs/10000] \approx 6.4 = 7$$

The digital filter method has been used to complete the filter response by using the following equation (1) :

$$H(e^{jw}) = \frac{1}{A(e^{jw})} = \frac{1}{a(1) + a(2)e^{-jw} + ... + a(n+1)e^{-jnw}} \quad (1)$$

Where a(1) and a(2) .. are the Linear predictor coefficient.

Step 2: Using equation (2), L_i vector where i= 0 to 44100 has been calculated using

$$L_i = 20 * log10 (abs(H) + eps) \quad (2)$$

The L_i Vector so obtained has been shown in Fig 3 for the musical instruments *Bansuri* and *Shehnai*
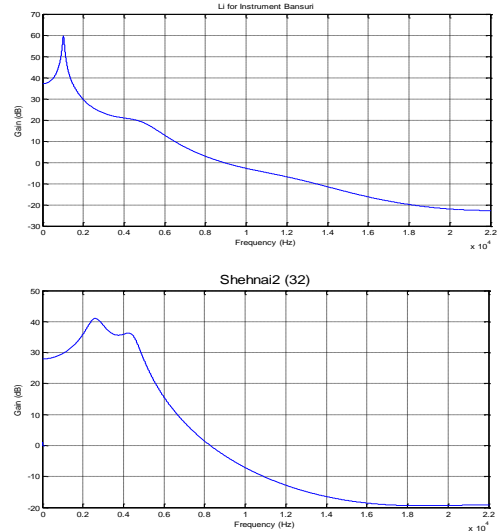

Fig 3: LPC Feature Vector L_i for i = 0 to 44100

Step 3: Rescaling of the features to unit range, typically a range between 0 and 1 is done by using Min-Max scaling. The normalize Vector P_i is represented as:

$$P_i = \frac{L_i - Lmin}{Lmax - Lmin}$$

Step 4: Without losing the information we grouped the signal with 1000 data points to obtain the P_i features set of 44 data points from average value of L_i as shown in Fig 4.
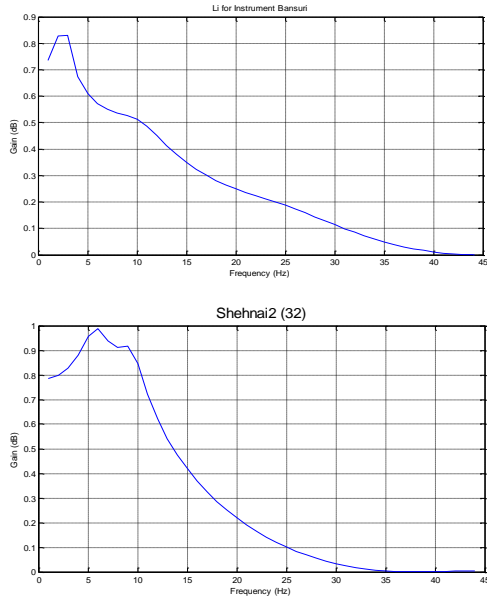
Fig 4: Reduced LPC Feature Vector P$_i$ for i = 1 to 44

Step 5: For the Final Feature set, we discarded the values close to 0 which was not effective, later we selected the odd features from the remaining P$_i$, minimizing the number of feature vector to the final size of P$_i$ where i= 1 to 15.

Step 6: Using the Final Feature Vector Set we tried to identify the musical instrument with the help of LDA Classification algorithm. LDA is the final processing step for Indian Musical Instruments Recognition task.

4.2 Dataset

For the evaluation of the current proposed system, we have selected nine Indian Musical Instruments. They are: *Bansuri (60), Shehnai (60), Dholki (50), Ghungaroo (60),* Triangle *(40), Ghatam (90), Manjira (60), Dhaf (90) and Morsing (70).* The figure in the bracket next to the musical instrument name indicates the total number of samples considered. The Digital Signal of each musical instrument was recorded at a sampling rate of 44.1 KHz and a bit depth of 16-bits.

5. RESULTS

As Seen in Fig 5, it has been very convenient to classify/group the nine musical instruments.
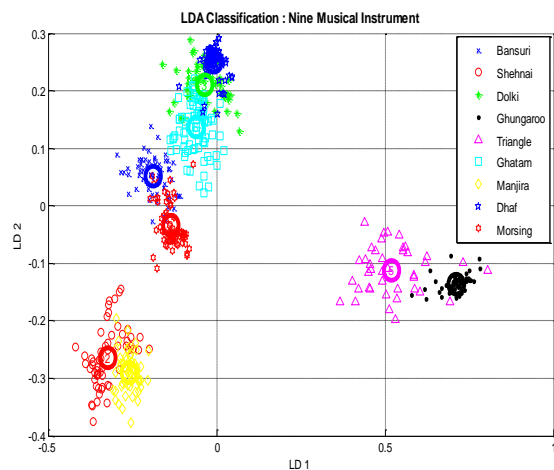

Fig. 5 Nine Musical Instrument LDA Classification

From the LDA features set, we also computed the center of data of each individual group of instruments samples using the equation as:

$$Ri = (Xci, Yci) = (\sum Xi / N , \sum Yi/ N)$$

where N – No. of Samples within each Instruments.

As shown in table 1, we computed the distance between the two center of Musical Instrument using the equation:

$$\Delta(Rij) = \sqrt{(Xci^2 - Xcj^2) + (Yci^2 - Ycj^2)}$$

Experimental results suggest that *Ghungaroo,* Triangle and *Ghatam* are relatively separable while *Dhaf* and *Dholki* resemble more closely.

Table 1: Relative Distance Matrix: Δ (Rij)

	a	b	c	d	e	f	g	h	i
a	0	0.34	0.22	0.92	0.73	0.15	0.35	0.27	0.1
b	0.34	0	0.56	1.04	0.86	0.48	0.07	0.6	0.3
c	0.22	0.56	0	0.82	0.64	0.08	0.55	0.05	0.26
d	0.92	1.04	0.82	0	0.19	0.82	0.98	0.81	0.85
e	0.73	0.86	0.64	0.19	0	0.63	0.8	0.64	0.66
f	0.15	0.48	0.08	0.82	0.63	0	0.47	0.12	0.19
g	0.35	0.07	0.55	0.98	0.8	0.47	0	0.59	0.28
h	0.27	0.6	0.05	0.81	0.64	0.12	0.59	0	0.31
i	0.1	0.3	0.26	0.85	0.66	0.19	0.28	0.31	0

a: Bansuri, b: Shehnai, c: Dholki, d: Ghungaroo, e: Triangle,
f: Ghatam, g: Manjira, h: Dhaf and *i: Morsing*

We split the dataset of 580 Samples of nine musical instrument in training set (66%) with 383 samples and testing data set (34%) having 197 samples. The confusion matrix obtained using the Best First Decision tree algorithms with the 10 fold cross validation is shown in table 2. It is observed that out of 197 samples tested we outperformed the recognition rate upto 93.04% (Correctly Classified Instances-184). The major error occurred in the detection of sound of *Dholki,* which was classified incorrectly as the *Dhaf.*

Table 2: Confusion Matrix using Best First Decision Tree

a	b	c	d	e	f	g	H	i	classified as
19	0	1	0	0	0	1	0	0	**a**
0	21	0	0	0	0	0	0	0	**b**
1	0	13	0	0	0	0	1	0	**c**
0	0	0	26	1	0	0	0	0	**d**
0	0	0	0	16	0	0	0	0	**e**
0	0	1	0	0	22	0	0	0	**f**
0	0	0	0	0	0	21	0	0	**g**
0	0	5	0	0	0	0	27	0	**h**
1	1	0	0	0	0	0	0	19	**i**

a: Bansuri, b: Shehnai, c: Dholki, d: Ghungaroo, e: Triangle,
f: Ghatam, g: Manjira, h: Dhaf and *i: Morsing*

6. CONCLUSION

The present research project has proposed an innovative method for Indian musical instrument recognition. This method is based on the features obtained from the modified Linear Predictive Coefficient approach and the classification method called Linear Discriminant Analysis. The present methods have obtained the Results that indicate 93.04% instruments are correctly classified by the Best First Decision Tree Algorithms. The proposed method outperformed the reference methods by a wide margin, indicating the modified LPC feature approach that provided a significant robustness improvement.

We propose to focus on comprehensive examination of spectral as well as temporal features in our research work. It would be interesting to examine analogical results when the complexity of these databases increase. Such an increase would become prominent when a large number of instruments are played and there will be several samples of individual instruments.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Glenn Eric, et al. "Hierarchical Parameterisation and classification for Musical Instrument Recognition" The 11th International Conference on Information Sciences, Signal Processing and their Applications.

[2] Eichhoff, Markus, and Claus Weihs. "Musical instrument recognition by high-level features." Challenges at the Interface of Data Analysis, Computer Science, and Optimization. Springer Berlin Heidelberg, 2012. 373-381.

[3] Eronen, Antti, and Anssi Klapuri. "Musical instrument recognition using cepstral coefficients and temporal features." Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on. Vol. 2. IEEE, 2000.

[4] Giannoulis, Dimitrios, and Anssi Klapuri "Musical instrument recognition in polyphonic audio using missing feature approach." Audio, Speech, and Language Processing, IEEE Transactions on 21.9 (2013): 1805-1817.

[5] Shah, Jashmin K., et al. "Robust voiced/unvoiced classification using novel features and gaussian mixture model." IEEE International Conference on Acoustics, Speech, and Signal Processing. 2004.

[6] Zlatintsi, Athanasia, and Petros Maragos. "Multiscale fractal analysis of musical instrument signals with application to recognition." Audio, Speech, and Language Processing, IEEE Transactions on 21.4 (2013): 737-748.

[7] Hai, Jiang, and Er Meng Joo. "Improved linear predictive coding method for speech recognition." Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on. Vol. 3. IEEE, 2003.

[8] Gunasekaran, S., and K. Revathy. "Fractal dimension analysis of audio signals for Indian musical instrument recognition." Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on. IEEE, 2008.

[9] Kim, Youngmoo E., and Brian Whitman. "Singer identification in popular music recordings using voice coding features." Proceedings of the 3rd International Conference on Music Information Retrieval. Vol. 13. 2002.

[10] Delac, Kresimir, Mislav Grgic, and Sonja Grgic. "Independent comparative study of PCA, ICA, and LDA on the FERET data set." International Journal of Imaging Systems and Technology 15.5 (2005): 252-260.

[11] Delac, Kresimir, Mislav Grgic, and Sonja Grgic. "A comparative study of PCA, ICA, and LDA." Proc. of the 5th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services. 2005.

[12] Toygar, Önsen, and Adnan Acan. "Face recognition using PCA, LDA and ICA approaches on colored images." Journal Of Electrical & Electronics Engineering 3.1 (2003): 735-743.

[13] Balakrishnama, Suresh, Aravind Ganapathiraju, and Joseph Picone. "Linear discriminant analysis for signal processing problems." Southeastcon'99. Proceedings. IEEE. IEEE, 1999.

[14] Ye, Jieping, Ravi Janardan, and Qi Li. "Two-dimensional linear discriminant analysis." Advances in neural information processing systems. 2004.

[15] Kohavi, Ron and George H.John, "Wrappers for feature subset selection" Artificial Intelligence 97.1 (1997) : 273-324.

[16] Hall, Mark,et al. "The WEKA data mining software: an update." ACM SIGKDD explorations newsletter 11.1(2009): 10-18.

[17] Shi, Haijian. Best-first decision tree learning. Thesis, The University of Waikato, 2007.