International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 31 • 2017

An Empirical Homophone Ambiguity Compression (HAC) from Involuntary Speech Identification

Papri Ghosh[a] Tejbanta Singh Chingtham[b] and Mrinal Kanti Ghose[b]

[a]Computer Science & Engineering, SIEM, MAKAUT, India

E-mail: papri.mss@gmail.com

[b]Computer Science & Engineering, SMIT, Sikkim, SMU

E-mail: chingtham@gmail.com, mkghose2000@yahoo.com

Abstract: An Involuntary Speech Identification (ISI) has been enlarged for constructing an optical text blueprint of the Speech Input for encroachment of Science and Technology. People are being upgraded to be enabling for indulgent the voice communication in an alternate way and pursue instruction using their optical skill. In the secluded communication, the optical skill becomes more influential than the listening skill and here ISI plays an imperative liability. During ISI, the system faces many ambiguities. The proposed research focuses on the Homophone Ambiguity and with the Involuntary learning it aims to improve it moderately. In the proposed Empirical Homophone Ambiguity Compression (HAC), a large amount of dataset is taken as homophones and Homophone Sets are being accumulated by Hierarchical Clustering Method (HCM). The proposed system corresponds with the user in case of Homophones and converts them to text layout with suitable credentials.

Keywords: *Human Computer Interaction (HCI), Homophones, Involuntary Speech Identification (ISI), Ambiguity, Machine Learning, Homophone Ambiguity Compression (HAC), Speech Identification System (SIS).*

1. INTRODUCTION

Human-computer interaction (HCI) or Man-Machine Interaction is a neighborhood of research which is well-liked because of it swiftness and ease of deployment [1-4]. It materialized in the early 1980s, initially as a branch of learning area in computer science espousal cognitive science and human concerned engineering. HCI has urbanized hastily and steady since last three decades, sketched courtesy of professionals from numerous other regulation and integrating divergent perception and approaches [5-6]. Up to a considerable level, the Man-Machine Interaction now combines a compilation of semi-autonomous countryside to discover and run through in human-centered informatics. However, the progressing amalgamation of different conceptions and move towards to science and practice in Man-Machine Interaction has produced a theatrical example of how unlike paradigms can be acquiescent and integrated in a vigorous and artistic rational project [7-9].

There are a group of transitional approach utilized for Man-Machine Interaction Design *i.e*. Cognitive Approach, Empirical Approach, Predictive Modeling Approach [10] etc. Computer System accepts several types of instructions and produces comparable outputs *i.e*. text, image, speech, texture, gesture etc. In recent times Speech Inputs are used massively especially in the smart phones and in the modern computers.

Regrettably, Researchers are facing few restrictions in the Involuntary Speech Identification. The homophone ambiguity is one of the major drawbacks in the existing Involuntary Speech Identification [12]. Homophones are set of two or more Similar Sounding Words with altered meanings or spelling [*i.e*. Eye/I, Right/Write/Wright, Beer/Bear etc.] [13-21]. From a stage of verdict, it can be sometime achievable to discriminate between two dissimilar homophones by analyzing the grammar, clauses and phrases. But in the word level it is very tricky to decide the exact homophone during Speech Identification. Few most advanced commercially existing Speech Identification software like SIRI [22], Dragon Speech [23], SAPI [24] etc. are still facing homophone ambiguity in their system. The ongoing research is focusing on resolving the homophone ambiguity problem from Involuntary Speech Identification.

Machine Learning has become one of the strongholds of information technology. As the technology have penetrated in the era of outsized data set, a deluge of data calls for an mechanized method of data investigation is compulsory [25-31]. The analysis of detecting pattern robotically and application of revealed patterns to predict some future data is known as Machine Learning. Machine Learning learns the struggle with uncertainty and solves them by relating some Empirical and Hypothetical Algorithm. In the Empirical learning method, training dataset of categorized data is prearranged. The system response dependent upon the given data and take the subsequent action. In hypothetical or unsupervised learning the task is to describe the secreted arrangement from an unlabelled data. Cluster Analysis is the best method in unsupervised learning. The research applies the Empirical Learning Mechanism in the Involuntary Speech Identification (ISI) for partial reduction the involvedness due to Homophone Ambiguity during recognition.

2. HOMOPHONES AMBIGUITY

Ambiguity is the chief anxiety in Involuntary Speech Identification (ISI) as the system is incapable to differentiate between two or more than two homophones during recognition in the word level. The probability of differentiating between the homophones in the sentence level with the help of the language learning and grammatical learning method is superior to identify within word level. The challenge of compliant the accurate homophone during the Speech Identification is received by the proposed research. In the proposed Speech Identification technique, the system inspects an English dictionary where the Homophone words are also accumulated. The system is not intelligent to differentiate between two or more homophones and cannot make a decision for which homophone the user is eying for. As a restriction, in most of the attempts, the system receives the initial homophone word among all its other homophones accumulated there in the system dictionary. Thus the remaining homophones are not been detected by the system. In the English dictionary the homophones are stored in an alphabetic order. Considering *Right, Rite, Wright* and *Write*; '*Right'* is being recognized by the ISI at most time. There are fewer amount of probability of receiving the alternate homophones inside the similar homophone set. For resolving the suspected difficulty, a homophone database is prepared from numerous resources and assembled them by Hierarchical Clustering Method [32]. The homophone database then joined and their redundancies are removed by using Microsoft Excel 2007 [33-34]. Total 1526 words were taken as homophones in the homophone database where 647 sets are there with homophones set. Within the homophone database, equivalent homophones are assigned a *set value*. Soon after the detection of a homophone by the *Speech Identification System (SIS)*, the system is presentation all the other homophones to the user those are belonging to the identical *Homophone Set*. By getting the entire alternatives, the user is to decide the suitable phrase for which he/she is eyeing for. Thus the probability of being accepted by the system has been increased. The sample database of homophones is shown in the Table. 1.

Table 1
Sample Database of Homophones

S. No.	*Contained Homophones*	*Total No of Members per set*
1.	Bity, Byte	2
2.	Cent, Scent, Sense, Sent	4
3.	Click, Clique	2
4.	Ran, Run	2
5.	To, Too, Two	3

3. HAC: SUPERVISED ALGORITHM

This section portrays the Empirical Homophone Ambiguity Compression (HAC) homophone ambiguity reduction from the Involuntary Speech Identification (ISI) System during detection.

3.1. Algorithm: HAC

Input : I_P (Speech I/P), H_{DB} (Homophone Dataset)

Output : T_C // Converted Text

Steps :

1. Start
2. $\hat{S} = \bar{I}(T_C(I_P))$ // Initial Converted Text
3. IF $\hat{S} \in H_{DB}$ THEN
4. Display $\hat{S}_{H[1..(n-1)]}$ //Homophone from $\hat{S}$
5. $T_C = H_{SEL}$ // Selected Homophone
6. END IF
7. Display T_C
8. REPEAT if required.
9. STOP

3.2. Flowchart of 'HAC'

The flowchart shows the flow of the algorithm in a pictorial representation. The Fig.1 illustrates the flowchart of the proposed Empirical Homophone Ambiguity Compression (HAC):

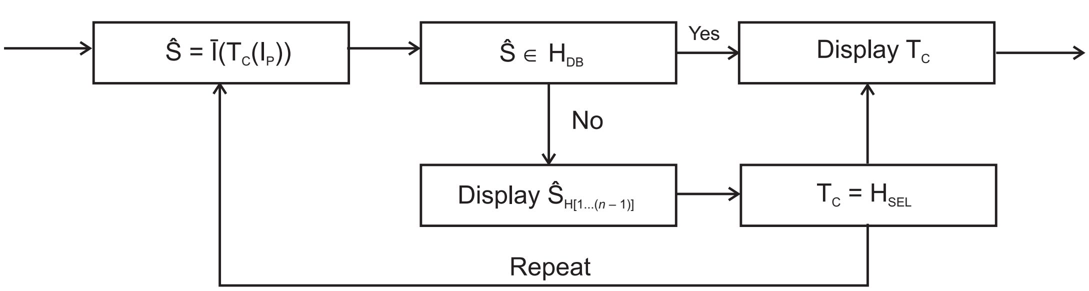

Figure 1: Flow Control of HAC

3.3. Over the flowchart, the method of being accepted of the précised homophone word is shown

Complexity Analysis of Empirical Homophone Ambiguity Compression (HAC)

Let $\Psi(\eta)$ be the total number of time essential for Empirical Homophone Ambiguity Compression (HAC). Let Ψ denotes max ($\Psi(\eta)$) over in all homophone set where η is the entire quantity of homophones per homophone set.

Lemma 1: To store a string β it is required a register with at least ɱ bits where ɱ is the number of total character in S.

***Proof*:** As string is a character array, the total no of bit compulsion β is proportional to the value of ɱ.

$$\beta \;\alpha\; ɱ \tag{1}$$

Lemma 2: If Homophone set does not contain β, then according to HAC it directly takes $\Psi(4)$ where Ψ denotes the worst time complexity.

***Proof* :** As HAC includes only 4 steps if it is not satisfying the IF-case in Step-3.

Lemma 3: The worst case complexity of HAC $\Psi(\eta) = O(\eta)$ where η is the total number of homophones within the homophone set.

***Proof* :** The Homophone set containing string β goes through the four steps within the IF-case. Let $\Psi(\eta)$ be the number of intervals allocated to a homophone set where the total number of homophones are η.

$$\Psi(\eta) = O(\eta) \tag{2}$$

3.4. Execution of HAC

In the proposed system, let us consider 'BYTE' as user given homophone that belongs to homophones set. The homophone of 'BYTE' is 'BITE'. The system recognizes 'BITE' instead of 'BYTE' as 'BITE' occurs first in the dictionary. After recognizing the 'BITE' if it is not user prefered, the system shows all other alternative. The user is allowed to choose from a list of homophone words if the correct one has not been automatically selected by opting for the corresponding number against the correct word. After that the system is showing the user prefered homophone word for which the user is looking for. When the user chooses a word which is not belongs to the homophone database, the system directly shows the word to the user.

4. ANALYSIS OF RESULT

The Empirical Homophone Ambiguity Compression (HAC) from Involuntary Speech Identification has been examined with twenty five male and female altered users. Several homophone words uttered by the users where only some go with the homophone database and some of them are not. The HAC is providing better effect in case of the homophonic words. As the proposed algorithm is not designed for the non homophonic words and it is presenting the similar result for them alike before applying the algorithm. The Fig.2 is showing the experimental status of the system before and after implementing the proposed algorithm on randomly taken three words which is not belonging to the homophone database.

An analysis of homophone words uttered by same set of users before and after applying HAC has been accomplished. Dependent upon the results, three categories of homophones identified within the database. The first category of homophones (*i.e.* "CENT") is taking place at the first position in the system dictionary among all its other homophone words and is recognized by the system before and after applying the proposed algorithm. The second type of homophone word (*i.e.* "Run") occurs after its first occurred homophone in the system dictionary; still it is being recognized by the system before applying the HAC algorithm. The third types of homophone word are (*i.e.* "Byte") occurred after its first occurred homophone in the system dictionary and never being accepted by the system. After applying HAC, the acceptance probability of second and third category of homophone words has been enhanced.

4.1. Probability Analysis of Acceptance of Homophones

After went through the entire homophone words from the homophone database and categorized them into three types, their acceptance rate has been analysed over HAC. In the homophone database total 1526 homophone word has been taken where total no of homophone set are 690. All the homophones are analyzed and tested before and after applying the Empirical Homophone Ambiguity Compression (HAC). Out of 1526 words 690 words are belonging to the first category. In rest of the homophone words, 562 are belonging to the third type of homophone words and 274 are from the second type of homophone words which may be recognized or maybe not. The test dataset can be increased more by adding more homophones in the homophone database. The chances of being accepted of the second categorized homophone words (*i.e.* 274) has been improvised and the third category (*i.e.* 562) words from the mentioned database has enhanced from zero to one. Fig.2 shows the analysis of homophonic words from the homophone database.

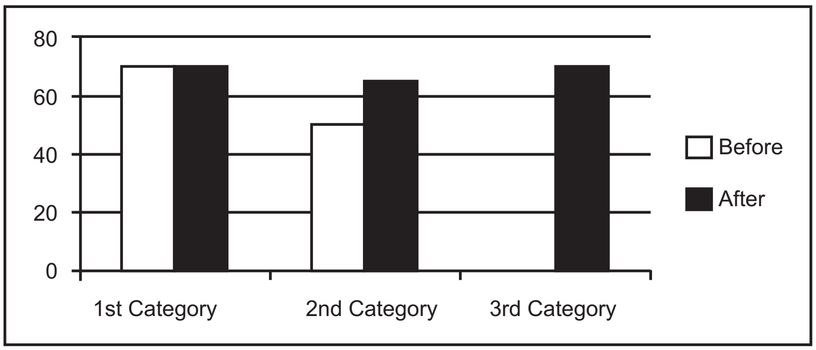

Figure 2: Experimental Status of Homophone Words before and after applying Empirical Homophone Ambiguity Compression (HAC)

Let us take δ as the homophone words from three different types, where δ_1, δ_2 & δ_3 denotes the first, second and third homophone set respectively. The following equations show the acceptance probability of homophone words from three different categories.

Before Applying Empirical Homophone Ambiguity Compression (HAC)

$$\Rightarrow \quad (\delta_1) = 1 \text{ [Probability of Sure Event]}$$

$$\Rightarrow \quad (\delta_2) = \frac{\gamma}{\acute{\eta}} \text{ [Probability of Partial Sure Event]}$$

$$\Rightarrow \quad (\delta_3) = 0 \text{ [Probability of Non-Sure Event]}$$

After Applying Empirical Homophone Ambiguity Compression (HAC)

$$\Rightarrow \quad (\delta_1) = 1 \text{ [Probability of Sure Event]}$$

$$\Rightarrow \quad (\delta_2) = \frac{\gamma + €}{\acute{\eta}} \text{ [Probability of Partial Sure Event]}$$

$$\Rightarrow \quad (\delta_3) = 1 \text{ [Probability of Non-Sure Event]}$$

Here $\frac{\gamma}{\acute{\eta}}$ denotes the initial probability of a homophone word of being accepted by the system where $\gamma, \acute{\eta} > 0$ & $\gamma < \acute{\eta}$ where $\acute{\eta}$ is a total number of homophones. $\frac{\gamma + €}{\acute{\eta}}$ denotes the increased probability of the δ_2 where $€ \geq 0$ and $\gamma + € < \acute{\eta}$.

The Speech Recognition System (SRS) contains all the barrier of ambiguities included noise, homophone, pronunciation etc. (assuming total n number of total ambiguities), then,

$$Ḁ_{\text{SR}} \;=\; \sum_{k=1}^{n} Ḁ_{k} \tag{3}$$

If ⇒ (SRS α $Ḁ_{\text{Homophone}}$) = 1, the HAC endow with consequence in the mentioned approach. Else the breach of 0 and $\text{Result}_{\text{SRS}}$ will be curtailed by other ambiguities. The proposed SRS embedded with HAC tested on a set of data that tested on Google Speech on the word level. In the sentence level, lexical or linguistic analysis helps SRS to recognize the correct homophones. For the mentioned SRS embedded with HAC,

$$\Rightarrow \qquad Ḁ(\text{Test}_{\text{Sentence}}) \;=\; 0 \tag{4}$$

So if the mentioned output is Ř, then

$$Ř \;\alpha\; Ḁ_{\text{Homophone}} \text{ iff } Ḁ_{\text{SR}} - Ḁ_{\text{Homophone}} \;=\; 0 \tag{5}$$

But practically, the following properties satisfies due to other Ḁ,

$$\Rightarrow \qquad Ḁ(\text{Result}_{\text{SRS}} \;\alpha\; Ḁ_{\text{Homophone}}) \;<\; 1 \tag{6}$$

5. FUTURE SCOPE

Considering a Finite Automata (FA) for the SRS $M_{SRS} = (Q, \Sigma, \delta, q_0, F)$ where $Q = \{q_0, q_1, q_2, q_f\}$ [q_0 = Initial SRS prior to Speech Input, q_1 = SRS following homophone input, q_2 = SRS if certain untrue homophones, q_f = SRS if accurately functioning, $\Sigma = \{h, h_n, h_c, h_{nc}, p_u\}$ [h_n = non-homophone, h_c = correct-homophone, h_{nc} = non-correct homophone, p_u = user-preference], F = {$q4$}, and δ defines in Table.2.

Table 2
Transition Function Table for M_{SRS}

State/Σ	*Output State*			
	h_n	h_c	h_{nc}	p_u
q_0	q_f	q_1	q_1	–
q_1	–	q_f	q_2	–
q_2	–	–	–	q_f
q_f	–	–	–	–

The transition system for M_{SRS} is given.

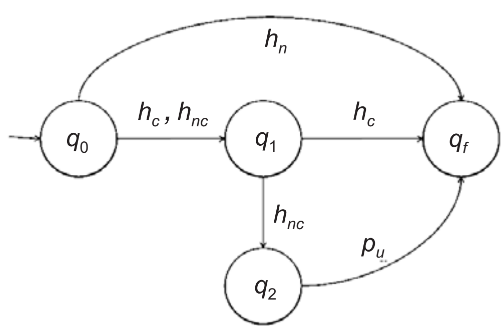

Figure 3: Transition System for M_{SRS}

From the above Finite Automata M_{SRS}, we get

$$R_E \;=\; h_n + (h_c + h_{nc})\,(h_c + h_{nc}\;.\;p_u) \tag{7}$$

Eliminating the non-homophone inputs from M_{SRS}, we get

$$R_E \;=\; (h_c + h_{nc})\,(h_c + h_{nc}\;.\;p_u) \tag{8}$$

If h_n is eliminated, then for all type of homophone inputs q_0 and q_1 can be considered as a same state and we may get the following automata $M1_{SRS}$

$$R_E = h_c + h_{nc} \cdot p_u \tag{9}$$

Let Finite Automata (FA) for *eq*. 9 is $M1_{SRS} = (Q1, \Sigma 1, \delta 1, q1_0, F1)$ where $Q1 = \{q_0, q_1, q_f\}$ [q_0 = Initial SRS prior to Speech Input, q_1 = SRS following incorrect homophone input, q_f = SRS if accurately functioning, $\Sigma 1 = \{h_c, h_{nc}, p_u\}$ [h_c = correct-homophone, h_{nc} = non-correct homophone, p_u = user-preference], $F1 = \{q_f\}$, and $\delta 1$ defines in Table.3.

Table 3
Transition Function Table for $M1_{SRS}$

State/ Σ	*Output State*		
	h_c	h_{nc}	p_u
q_0	q_f	q_1	–
q_1	–	–	q_f
q_f	–	–	–

The transition system for $M1_{SRS}$ is given.

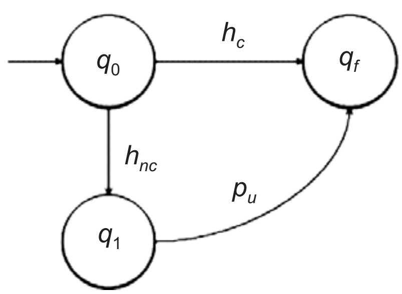

Figure 4: Transition System for $M1_{SRS}$

The propose research aims to formulate

$$h_c = h_{nc} \cdot p_u \tag{10}$$

So,

$$h_{nc} = \{h_c \mid p_u = \Lambda\} \tag{11}$$

The future effort of this research is to focus to formulate the subsequent Eq. 12,

$$p_u \rightarrow \Lambda \tag{12}$$

Though the precision of eq. 13 establish a new prospect of the enhancement of the contemporary supervised research into an unsupervised research.

$$\Rightarrow \quad (\Rightarrow (h_c = h_{nc})\ \alpha$$

$$\Rightarrow \quad (p_u \rightarrow \Lambda)) < 1 \tag{13}$$

6. CONCLUSION

The research only focuses on the empirical study where the system deals with the well-known set of data by which applying a specified set of trained instruction the system can accomplish an expected set of outputs. An immense number of Homophones were assembled from quite a lot of sources and then they were combined with no redundancies. The Empirical Homophone Ambiguity Compression (HAC) was applied on preferred homophone words. The entire comparative results have been modelled in this paper. However, at the early stage of the system, there were some reconsideration and refinements that could be made to achieve more appropriate result pertaining some empirical approach. The research is further extended with a hypothetical approach where the task is to describe the hidden configuration from unlabelled facts. There are many avenues that can be discovered with the further research. One is the analogy of this work with empirical learning strategies, partial augmentation gained in Speech Recognition System (SRS) during recognition.

REFERENCES

[1] B. A. Myers. "A Brief History of Human Computer Interaction Technology." ACM interactions. Vol. 5, No. 2, pp. 44-54, March, 1998.

[2] M. Fetaji, S. Loskoska, B. Fetaji, M. Ebibi. "Investing Human Computer Interaction Issues in Designing Efficient Virtual Learning Envioronment". BCI, Sofia, Bulgaria, 2007.

[3] K. P. Tripathi, "A Study of Interactivity in Human Computer Interaction", International Journal of Computer Applications, ISSN: 0975 – 8887, Vol. 16, Iss. 6, Feb 2011.

[4] P. Sharma, N. Malik, N. Akhtar, Rahul, "Human Computer Interaction", International Journal of Advanced Research in IT and Engineering (IJARITE), ISSN: 2278-6244, Nov 2012.

[5] S. Agravat, G. Pandey, "Human Computer Interaction using Computer Vision," IJCSMR (ISSN 2278-733X), Vol. 2, Issue 1, Jan 2013.

[6] "ACM SIGCHI Curricula for Human-Computer Interaction", Definition of HCI, [Available: http://old.sigchi.org/cdg/cdg2.html] [Access Date: May 21, 2014].

[7] "Human Computer Interaction (HCI)", [Available: http://www. webopedia.com/TERM/H/HCI.html] [Access Date: May 21, 2014].

[8] "Human-computer Interaction", [Available: https://en.wikipedia. org/wiki/Human–computer_interaction] [Access Date: May 21, 2014].

[9] P. Ghosh, C. T. Singh, M. K. Ghose, "A Review of Modern HCI Challenges for Betterment of Differently Abled People", IJCIR, Vol. 9, Issue. 2, pp 79-88, Nov 2013.

[10] "HCI Design Approaches" [Available at: http://www.usabilityfirst. com/usability-methods/hci-design-approaches/] [Access Date: 6th June, 2016]

[11] S. Gujral, M. Tuteja, B. Kaur, "Various Issues in Computerized Speech Recognition Systems" , IJERGS, Vol.2, Issue. 4, June-July 2014, ISSN 2091-2730.

[12] M. Forsberg, "Why is Speech Recognition Difficult?", [Available: http://www.speech.kth.se/~rolf/gslt_papers/MarkusForsberg.pdf] [Access Date: Dec 15, 2015].

[13] "List of Homophones" [Available: http://www.allaboutlearningpress. com/list-of-homophones/] [Access Date: July 30, 2014].

[14] "Homophone" [Available: http://en.wikipedia.org/wiki/Homophone] [Access Date: July 30, 2014].

[15] "Homophones", [Available At: http://www.singularis.ltd.uk/bifroest /misc/homophones-list.html] [Access Date: Dec 14, 2015].

[16] "Homophones", [Available At: http://www.allaboutlearningpress.com/ homophones/] [Access Date: Dec 14, 2015].

[17] "Common Homophones List", [Available At: https://www.englishclub .com/pronunciation/homophones-list.htm] [Access Date: Dec 14, 2015].

[18] "Homophones Word List", [Available At: http://www.lccc. edu/sites /www.lccc.edu/files/images/LCCC%20homophones2.pdf] [Access Date: Dec 14, 2015].

[19] "Homophones Examples with Definitions", "[Available At: http://www .grammarinenglish.com/homophones/]] [Access Date: Dec 14, 2015].

[20] "Homophones by Letter", [Available At: http://www. homophone.com/browse] [Access Date: Dec 14, 2015].

[21] "The Megalist of Homophones", [Available At: http://fivejs.com/ homeschool-downloads/homophones-list/] [Access Date: Dec 14, 2015].

[22] B. Johnson, "How SIRI Works: What are SIRI's Limitations?" [Available: http://electronics.howstuffworks.com/gadgets/high-tech-gadgets/siri4.htm] [Access Date: Mar 14, 2015].

[23] NUANCE, "Dragon Naturally Speaking" [Available: https://www. bolton.ac.uk/ITSupportServices/IT-Publications/Workbooks/ Dragon - 115-user-guide.pdf] [Access Date: Mar 14, 2015].

[24] "Speech Application Program Interface (SAPI) – Part of the Microsoft Glossary" [Available At: http://whatis.techtarget.com/definition/Speech-Application-Program-Interface-SAPI] [Access Date: 21st May, 2016]

[25] H. Duame, "A Course in Machine Learning", [Available at: http://ciml.info/dl/v0_8/ciml-v0_8-all.pdf] [Access Date: 21st May, 2016]

[26] C. M. Bishop, "Pattern Recognition & Machine Learning", Springer ISBN: 0387310738, 1st Oct, 2007.

[27] A. Smola, S. V. N. Vishwanathan, "Introduction to Machine Learning", [Available at: http://alex.smola.org/drafts/thebook.pdf] [Access Date: 18th May, 2016]

[28] E. Alpaydin, "Introduction to Machine Learning", [Available at: http://cs.du.edu/~mitchell/mario_books/Introduction_to_Machine_Learning_-_2e_-_Ethem_Alpaydin.pdf] [Access Date: 19th May, 2016]

[29] K. P. Murphy, "Machine Learning: A Probabilistic Perspective", The MIT Press, Cambridge, London, England.

[30] T. M. Mitchell, "Machine Learning", McGraw Hill, ISBN: 0070428077.

[31] S. S. Shwartz, S. B. David, "Understanding Machine Learning: From Theory of Algorithms", Cambridge University Press, 2014. [Available At: http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/ understanding-machine-learning-theory-algorithms.pdf] [Access date: 18th May, 2016]

[32] L. Rokach, O. Maimon, "Clustering Methods", "Data Mining and Knowledge Discovery Handbook", Springer, ISBN.8184891393. [Available at: https://www.cs.swarthmore.edu/~meeden/cs63/s16/ reading/Clustering.pdf] [Access Date: 22nd May, 2016]

[33] "Removing Duplicates in Excel" [Available: http://www.mrexcel .com/tip138.shtml] [Access Date: Mar 01, 2015].

[34] "Filter for Unique Values or Remove Duplicates Values" [Available: https://support.office.com/en-us/article/Filter-for-unique-values-or-remove-duplicate-values-ccf664b0-81d6-449b-bbe1-8daaec1e83c2] [Access Date: Mar 01, 2015].