Multi-stage Children Story Speech Synthesis for Hindi

Harikrishna D M, Gurunath Reddy M, K. Sreenivasa Rao
School of Information Technology
Indian Institute of Technology Kharagpur, Kharagpur - 721302, India
Email: {harikrishna.dm, mgurunathreddy}@sit.iitkgp.ernet.in, ksrao@iitkgp.ac.in

***Abstract*—In this paper, we propose a multi-stage children story speech synthesis system for Hindi language. The proposed system performs the following tasks: (i) classification of stories into different genres based on text, (ii) prediction of emotion from story text, (iii) deriving prosody rules (modification factors) specific to emotions and story genres and (iv) synthesis of story speech using mark-up language and prosody modification factors. Keyword and part-of-speech (POS) features are used for story-genre classification and emotion prediction. The prosody modification factors are derived carefully by analyzing the perceptual differences between synthesized neutral speech utterances and their respective utterances narrated by a storyteller. The story is synthesized by the festival based concatenative speech synthesizer with annotated story in the form of SABLE mark-up language. The quality and naturalness of the synthesized story speech is evaluated using subjective tests.**

***K*eywords —** Story speech synthesis, Story-genre classification, Story-genre-specific emotions, SABLE mark-up, Children story synthesis

I. Introduction

Children stories can be synthesized in many ways like in-domain unit selection, post-processed domain-independent unit selection, in-domain statistical parametric synthesis, model-adapted synthesis etc. Out of this, two most widely used approaches are : (i) in-domain unit selection i.e. by explicitly building concatenative text-to-speech (TTS) systems from the story speech corpus narrated by professional artists and (ii) post-processed domain-independent unit selection i.e. by post processing the synthesized speech from neutral TTS system. Though the former approach is easier to synthesize story, it requires collecting hours of story speech corpus from professional artists which is an expensive and laborious task. In the latter approach, prosody modification factors are derived from a small set of story corpus collected from professional artist and then neutrally synthesized speech is prosody modified using the derived modification factors. Recently, syllable-based unit selection neutral TTS systems were developed in 13 Indian languages by [1]. The goal of this work is to synthesize children story speech from neutral TTS systems for Hindi language.

In previous works, emotions in children stories were analyzed for child-directed expressive text-to-speech synthesis. In [2], a perceptual study on emotions based on expressive spoken corpus of children's stories was conducted and claimed that collaboration of semantic and prosodic cues helps in expressing emotional content in a story. For an emotional TTS synthesis, prediction of emotions from text passage is an important task. In [3], machine learning based model for text based emotion prediction was proposed. A multi-step system for analyzing children stories was proposed in [4]. Their system performed tasks like character identification; identification of attributes like character age, gender; emotion analysis of the quoted text. In order to synthesize either basic emotions and its variants or children stories with emotions, a post-processing approach is followed on a synthesized neutral speech. In [5], storytelling style speech is synthesized from a neutral concatenative speech synthesizer using prosody rule-set. Post-processing of synthesized neutral speech involves modifying the prosody parameters of the speech using available prosody modification tools. Though the synthesized speech is highly intelligible, it lacks naturalness and is prone to perceptual distortions.

To overcome these limitations, we are proposing a concatinative based story speech synthesizer. The synthesizer selects the natural speech units based on the prosody modification factors provided along with the input story text. While analyzing the stories, it is observed that stories can be broadly classified into three genres namely fable, folk-tale and legend. Hence we also derived separate prosody modification factors for the basic emotions with respect to different story genres.

The hypothesis of this work is that there are important interactions between emotion and story genre which need to be handled in prosody generation. We tried to judge this hypothesis by the subjective evaluation where genre-specific systems are compared with general emotion ones. The scope of the present work includes, automatic classification of the given input story text into one of the three genres. Then automatically predicting the emotions present in the classified story into one of the basic emotion categories. Deriving the prosody modification factors by analyzing the stories narrated by narrator and the corresponding neutrally synthesized story speech. Finally, synthesizing the children story from the given input story text by automatically converting it into SABLE mark-up format.

The rest of the paper is organised as follows: Section II gives an overview of the proposed framework. Details of story classification are described in Section II-A. Prediction of emotion for story synthesis is explained in Section II-B. Section II-C explains the rules for deriving prosody modification factors and story synthesis using SABLE mark-up language. Evaluation procedures are discussed in Section III. Section IV presents the experimental results and discussions. Conclusions and scope for future work are discussed in Section V.

II. Proposed Framework

Flow diagram of the proposed multi-stage story speech synthesis system is shown in Figure 1. Initially, the given story text is classified into one of the three genres namely fable, folk-tale and legend. Then emotions are predicted from the classified story. Finally, the story speech is synthesized by story specific prosody rule-set.

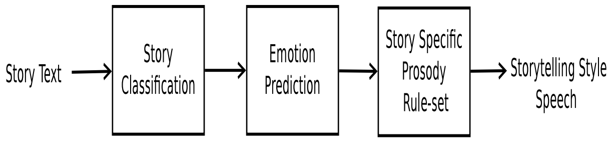

Fig. 1. Proposed framework for multi-stage story synthesis

A. Story Classification

Hindi story corpus consists of 300 short stories collected from blogs, Panchatantra and Akbar-Birbal books. The current story corpus has three different story genres namely, fable, folk-tale and legend. Fable is a short animal tale. Folk-tale is a popular story which is passed on in spoken form from one generation to the next. Legend is a semi-true story carrying important meaning or symbolism for the culture in which it originates. Details of the Hindi story corpus are presented in Table I.

TABLE I. Details of Hindi story corpus

Category	# Stories	# Words	# Unique Words
Fable	100	50344	3313
Folk-tale	100	46900	3710
Legendary	100	35991	2963

Corpus is cleaned as a part of pre-processing. Multiple white spaces are stripped and punctuation marks, special symbols and numbers are removed. Furthermore, Hindi Shallow Parser[1] is used for lemmatization to convert each word into its root word (base form). Lemmatization is an important task as it decreases the feature vector dimension which leads to better representation of document. Stopwords are the most common words which do not contain important significance in a language. There is no single standard stopword list and it is relative to the domain of the document. In this work, we listed top 500 most frequent words. Based on the semantic content of the word, a list of 170 stopwords was prepared and used in this work. Hindi story classification is treated as supervised machine-learning problem, where stories are projected into Vector Space Model (VSM) which uses words as features. Linguistic based features like density of POS (PD) and different weighting schemes like term frequency (TF) and term frequency inverse document frequency (TFIDF) are explored. Different combinations of features vectors are considered for evaluation. R statistical programming language is used for feature extraction [6].

- **Term Frequency (TF)**: Frequency of terms in a document are calculated. Importance of a word within a story genre is given by TF measure.
- **Term Frequency Inverse Document Frequency (TFIDF)**: For a term, weight is assigned as product of TF and IDF. IDF is calculated as

$$idf(t_i) = log\frac{N}{n_i}$$

  where N is the total number of stories and n_i is the number of stories in corpus that contains word t_i. Importance of a word across story genre is given by TFIDF measure.
- **POS Density (PD)**: Relevance of the POS tags namely Noun (NN), Proper Noun (NNP), Spatial and Temporal Nouns (NST), Pronoun (PRP), Adjective (JJ) and Quantifiers (QF) with respect to Indian languages are explained in shallow parser manual[2]. For each document, PD is used as feature vector. It is calculated as

$$PD = \sum_{p \in P} \frac{count(p)}{Total\ words\ in\ document}$$

  where $P = \{NN,\ NNP,\ NST,\ PRP,\ JJ,\ QF\}$

B. Emotion Prediction

In this work we have considered four emotions: happy, sad, anger and fear. Details of emotional sentences present in the story speech corpus are given in Table II. Emotion annotation task is hard since there is no clear definition for emotion and non-emotion. Four annotators were instructed regarding the annotation task. Each annotator labels the sentences with one of the four emotions mentioned above based on mood, intensity and semantics of the sentence. Sentence will be labelled as neutral if the annotator feels that there is absence any of the above mentioned emotions. Measure of agreement among multi-annotators are discussed in [7]. Fleiss Kappa (κ) is a statistical measure of inter-annotator agreement and is computed using online tool [8]. Fleiss Kappa (κ) = 0.774 is observed which is considered to be substantial agreement.

TABLE II. Details of emotional sentences

Emotion	# Sentences
Happy	214
Sad	232
Anger	108
Fear	76

Emotion prediction task can be considered as multi-class classification problem. Given a sentence, we have to classify it into one of the emotions mentioned above. For this task, a list of emotional words/phrases is prepared from the story corpus. It is observed that: (i) most of the emotional words are nouns, adjectives and adverbs and (ii) adjective-adverb pair and intensifier emphasizes on emotions. Hence POS and Sentence level features are considered for predicting emotions and are mentioned below. Feature extraction code is written in R.

- Count of emotional words/phrases in a sentence
- Count of adjective-adverb pairs in a sentence

[1]http://ltrc.iiit.ac.in/showfile.php?filename=downloads/shallow_parser.php

[2]http://ltrc.iiit.ac.in/tr031/posguidelines.pdf

- Count of intensifiers in a sentence
- Count of nouns in a sentence
- Count of adjectives in a sentence
- Count of adverbs in a sentence

C. Prosody Rules Derivation and Synthesis

A Zero frequency filtering (ZFF) based method capable of capturing the rapid variations of source parameters is presented in [9]. Hence, in this work, to capture the dynamically varying prosody parameters of story speech, the modified ZFF is used to derive the prosody parameters. The method uses ZFF signal derived from speech to obtain the instants of Glottal closures and the strength of excitation at the epoch locations.

ZFF signal is obtained as follows:

- Th input speech signal x[n] is passed through a cascade of zero frequency resonators given by

$$y_0[n] = -\sum_{k=1}^{4} a_k y_0[n-k] + x[n] \tag{1}$$

  where $a1 = 4$, $a2 = -6$, $a3 = 4$, and $a4 = -1$
- The trend in $y_0[n]$ is removed by subtracting the mean computed over a window at each sample. The resulting signal y[n] is the ZFF signal, given by

$$y[n] = y_0[n] - \frac{1}{2N+1} \sum_{m=-N}^{N} y_0[n+m] \tag{2}$$

  where $(2N + 1)$ is the size of the window, which is in the range of 1 to 2 times the average pitch period of the speaker.
- This method does not captures the rapid variations of F0 in the story. To capture the rapid variations of F0, the method is modified as in [10] using the following steps to derive the epochs and their strength of excitation (SoE) from the ZFF signal.
- Speech signal is passed through the Zero-frequency resonator with a widow length of 4 ms for trend removal. The energy of ZFF signal will be relatively high for voiced regions and low for unvoiced and silence regions.
- Hilbert envelope of the ZFF signal is computed. Envelope is smoothed by using a 10 ms running average filter. Smoothed envelope is threshold by 1% of the maximum sample value to get voiced and unvoiced segments.
- For each voiced segment, window length for trend removal is obtained by location of maximum peak in the autocorrelation function of that segment. Each voiced segment is filtered separately with corresponding window length.
- The epoch locations are obtained from the positive zero crossings of the final filtered signal, strength of excitation (intensity) is obtained by the slope of ZFF signal at epoch locations.
- Epoch interval plot is obtained by successive epoch location differences. Finally, pitch contour is obtained by the inverse of the epoch interval plot.
- The pitch and intensity contour plots for all synthesized neutral and their corresponding story emotive phrases are obtained. From the plots, prosody parameters such as pitch range (PR), pitch base (PB) and average intensity (Intensity) are calculated. The mean prosody parameters for all emotional phrases viz. anger, happy, sad and fear are obtained. The prosody modification factors are obtained by ratio of mean prosody parameter of the recorded story emotion to the mean neutral synthesized prosody parameter. For instance, the intensity modification factor of +30% indicates that the mean intensity of neutral speech has to be increased by 30%.
- Similarly the mean speaking rate (Rate) of a phrase is obtained by the ratio of duration of emotional phrase to that of neutral synthesized phrase. For instance, the mean speaking rate (Rate) of -25% indicates that the duration of the synthesized speech has to be reduced by 25%.

A small set of story corpus comprising of 30 stories were recorded by a female professional artist for deriving prosody modification factors. Out of 30, 15 stories were general stories i.e. stories which do not belong to fable, folk-tale or legend and the rest 15 stories comprised of 5 stories of fable, folk-tale and legend. These 30 stories are synthesized using Hindi neutral TTS system[1]. The prosody modification factors are derived carefully by analyzing the perceptual differences between synthesized neutral speech utterances and their respective utterances narrated by a storyteller. The prosody modification parameters derived at phrase level for emotions of generalized stories and emotions specific to story genre are shown in Tables III and IV respectively. From Table IV, it can be observed that the prosody modification factors indeed vary across story genres. In this work, SABLE mark-up language is used to synthesize story [11]. SABLE mark-up language supports many user-controlled tags to improve the naturalness of the synthesized speech from festival speech synthesizer. A subset of prosody tags such as pitch base, pitch range, speaking rate and intensity are used to synthesize story. After automatically annotating the classified story phrases into different emotion categories, it is converted into SABLE mark-up format using prosody tags. A sample template for fable story synthesis is shown in Figure 2. Now the story speech is synthesized by mark-up from the festival TTS system.

TABLE III. PROSODY MODIFICATION FACTORS FOR EMOTIONS OF GENERALIZED STORIES

Emotions	Intensity (%)	Rate (%)	PR (%)	PB (%)
Anger	+30	+30	+112	+20
Happy	+20	+20	+150	+30
Sad	-10	-22	-30	-15
Fear	+10	-10	+43	+20

TABLE IV. PROSODY MODIFICATION FACTORS FOR EMOTIONS SPECIFIC TO STORY GENRES

Story genres	Emotions	Intensity (%)	Rate (%)	PR (%)	PB (%)
Fable	Anger	+35	+30	+115	+20
	Happy	+20	+25	+150	+30
	Sad	-10	-20	-25	-15
	Fear	+10	-10	+40	+15
Folk-tale	Anger	+38	+35	+120	+25
	Happy	+20	+30	+160	+30
	Sad	-15	-18	-25	-20
	Fear	+15	-15	+50	+20
Legend	Anger	+40	+40	+125	+30
	Happy	+25	+35	+165	+35
	Sad	-20	-20	-25	-25
	Fear	+20	-10	+55	+25

SABLE mark-up: template for fable story synthesis

```
<SABLE>
    <FABLE>
        <Emotion Name="Happy">
        <Pitch Range = "+150">
        <VOLUME LEVEL = "+20"/>
        "Happy Phrase/Sentence"
        </Emotion>
        <Emotion Name="Sad">
        <Pitch Range = "-25">
        <VOLUME LEVEL = "-10"/>
        "Sad Phrase/Sentence"
        </Emotion>
        <Emotion Name="Anger">
        <Pitch Range = "+115">
        <VOLUME LEVEL = "+35"/>
        "Anger Phrase/Sentence"
        </Emotion>
        <Emotion Name="Fear">
        <Pitch Range = "+40">
        <VOLUME LEVEL = "+10"/>
        "Fear Phrase/Sentence"
        </Emotion>
    </FABLE>
</SABLE>
```

Fig. 2. Sample template for fable story

III. EVALUATION

In this work, we used WEKA [12] as a framework combined with LibSVM [13] for story genre classification and emotion prediction. For story genre classification, five different combinations of features are used: PD, TF, TFIDF, TF + PD, TFIDF + PD. Sentence level features as described in Section II-B are used for emotion prediction. Performance of story classification and emotion prediction are evaluated using three classifiers: Naive Bayes (NB), K-Nearest Neighbour (KNN) and Support Vector Machine (SVM). Classifier performance is evaluated using 10-fold cross validation. For KNN, Nine nearest neighbours are used i.e. k=9. For SVM, linear kernel is used with other default settings in WEKA. Results are evaluated using Precision (P), Recall (R) and F-measure (F).

For subjective evaluation, we considered 20 sentences comprising of 5 sentences of four emotions. Those 20 sentences are synthesized using two methods: (i) generalized prosody modification factors and (ii) story genre specific prosody modification factors. Mean opinion scores (MOS) and preference tests measures are used for subjective evaluation of synthesized story speech. The subjects evaluated the quality of speech in terms of naturalness and story style on a five-point scale (1: very poor, 2: poor, 3: fair, 4: good and 5: excellent). In preference tests, the subjects are asked to give preference between pair of story speech utterance synthesized using generalised prosody modification factors and story genre specific prosody modification factors. Subjects are used to select the better among the two synthesized speech utterance or they can prefer both. To avoid bias, speech files are played to the subjects in random sequence. For proper assessment of speech signals, subjects with speech knowledge are selected for listening tests. Listening tests are conducted with 10 research scholars in the age group of 23-30 years in the laboratory environment by playing the speech signals through headphones.

IV. RESULTS

Tables V and VI represents the classifier performances for story classification and emotion prediction task respectively. In case of story classification, it is observed that combining POS and keyword based features do not improve significantly the classifier performance. Due to high dimensionality of TF and TFIDF, adding PD features to TF or TFIDF shows little improvement in Precision, Recall and F-measures. TF + PD features has the highest F-measure compared to other features. In case of emotion prediction, highest F-measure is observed for sad emotion using SVM classifier. Among the classifiers, SVM models outperformed the other models.

Table VII show the average MOS for speech utterance synthesized using generalised prosody modification factors and story genre specific prosody modification factors. Table VIII show the results of preference test. For preference test, the subjects preferred speech utterances synthesized using story genre specific prosody modification factors in 80% of cases, 10% preferred speech utterances synthesized using generalised prosody modification factors and rest 10% gave equal preferences. The results from preference test indicates that the quality of synthesized speech using story genre specific prosody modification factors is significantly better than the generalised prosody modification factors.

V. CONCLUSIONS

This paper proposes a cascade of story classification, emotion prediction and prosody rule-set as a framework for storytelling speech synthesis. Different approaches that range from text processing to prosody rules derivation and incorporation of these rules for synthesizing story speech are employed. Keyword and part-of-speech (POS) features are used for story-genre classification and emotion prediction. The prosody modification factors are derived carefully by analyzing the perceptual differences between synthesized neutral speech utterances and their respective utterances narrated by a storyteller. The story is synthesized by the festival based concatenative speech synthesizer with annotated story in the form of SABLE mark-up language. For story classification, combining POS and keyword based features did not improve significantly the

TABLE V. STORY CLASSIFICATION PERFORMANCE

Story genres	Features	NB			KNN			SVM		
		P	R	F	P	R	F	P	R	F
Fable	PD	0.47	0.39	0.42	0.46	0.66	0.54	0.47	0.62	0.53
	TF	0.74	**0.83**	0.78	0.52	0.7	0.59	0.91	0.42	0.57
	TFIDF	0.58	0.71	0.64	0.53	0.57	0.55	0.85	0.41	0.55
	TF + PD	0.75	**0.83**	**0.79**	0.53	0.71	0.61	**0.92**	0.43	0.59
	TFIDF + PD	0.59	0.71	0.65	0.54	0.6	0.57	0.86	0.42	0.56
Folk-tale	PD	0.36	0.48	0.41	0.37	0.38	0.37	0.34	0.46	0.39
	TF	0.62	0.53	0.57	0.52	0.74	0.61	0.58	0.85	0.69
	TFIDF	0.43	0.47	0.45	0.51	0.58	0.54	0.45	0.61	0.52
	TF + PD	**0.64**	0.54	0.58	0.55	0.74	0.63	0.59	**0.86**	**0.7**
	TFIDF + PD	0.42	0.47	0.44	0.5	0.68	0.58	0.45	0.61	0.52
Legend	PD	0.54	0.46	0.49	0.6	0.34	0.43	0.67	0.23	0.34
	TF	0.76	0.79	0.77	0.51	0.86	0.64	0.86	0.94	0.9
	TFIDF	0.71	0.5	0.58	0.41	0.8	0.54	0.58	0.69	0.63
	TF + PD	0.77	0.81	0.79	0.57	0.75	0.65	**0.87**	**0.95**	**0.91**
	TFIDF + PD	0.71	0.49	0.58	0.44	0.77	0.56	0.58	0.69	0.63

TABLE VI. EMOTION PREDICTION PERFORMANCE

Emotions	NB			KNN			SVM		
	P	R	F	P	R	F	P	R	F
Happy	0.49	0.46	0.47	0.59	0.58	0.58	0.72	0.73	0.72
Sad	0.52	0.65	**0.58**	0.54	0.69	**0.6**	0.7	0.82	**0.75**
Anger	0.53	0.37	0.44	0.54	0.34	0.42	0.63	0.5	0.55
Fear	0.46	0.4	0.43	0.49	0.34	0.41	0.49	0.39	0.43

classifier performance. However we can conclude that TF + PD feature combination is the best choice for story classification. Emotion prediction using SVM and the proposed features show interesting performances. For subjective evaluation, there is improvement of average MOS for story genre specific prosody modification compared to speech utterance synthesized using generalised prosody modification factors. From preference test, it can be inferred that the quality of synthesized speech using story genre specific prosody modification factors is better than the generalised prosody modification factors.

Future work should include a study on variation of perceptual differences for emotions and neutral sentences. A detailed evaluation with more stories has to be done. Apart from Hindi, the current study can be extended to other Indian languages.

ACKNOWLEDGMENT

We are thankful to the annotators and subjects who volunteered for listening tests. The authors would like to thank Department of Information Technology, Govt. of India for supporting this research work, *Development of Text-to-Speech synthesis for Indian Languages Phase II*, Ref. no. 11(7)/2011HCC(TDIL).

TABLE VII. MEAN OPINION SCORES FOR STORY SYNTHESIS USING DIFFERENT METHODS

Method	MOS
Generalized prosody modification factor	2.9
Story genre specific prosody modification factor	**3.1**

TABLE VIII. PREFERENCE TESTS FOR STORY SYNTHESIS USING DIFFERENT METHODS

Method	Preference (%)
Generalized prosody modification factor	10
Story genre specific prosody modification factor	**80**
Both	10

REFERENCES

[1] H. A. Patil, T. B. Patel, N. J. Shah, H. B. Sailor, R. Krishnan, G. Kasthuri, T. Nagarajan, L. Christina, N. Kumar, V. Raghavendra *et al.*, "A syllable-based framework for unit selection synthesis in 13 Indian languages," in *Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)*. pp. 1–8, 2013.

[2] C. O. Alm and R. Sproat, "Perceptions of emotions in expressive storytelling." in *INTERSPEECH*, pp. 533–536, 2005.

[3] C. O. Alm, D. Roth, and R. Sproat, "Emotions from text: Machine learning for text-based emotion prediction," in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, pp. 579–586, 2005.

[4] E. Iosif and T. Mishra, "From speaker identification to affective analysis: A multi-step system for analyzing children stories," *EACL*, pp. 40–49, 2014.

[5] P. Sarkar, A. Haque, A. Dutta, G. Reddy, M. Harikrishna, P. Dhara, R. Verma, P. Narendra, B. Sunil, J. Yadav, and K. Rao, "Designing prosody rule-set for converting neutral TTS speech to storytelling style speech for Indian languages: Bengali, Hindi and Telugu," in *Seventh International Conference on Contemporary Computing (IC3)*, pp. 473–477, Aug 2014.

[6] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org

[7] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological Bulletin*, vol. 76, no. 5, p. 378, 1971.

[8] J. Geertzen, " Inter-rater agreement with multiple raters and variables," https://mlnl.net/jg/software/ira/, 2012, Retrieved June 10, 2015.

[9] K. S. R. Murty and B. Yegnanarayana, "Epoch extraction from speech signals," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1602–1613, 2008.

[10] S. A. Thati, S. Kumar, and B. Yegnanarayana, "Synthesis of laughter by modifying excitation characteristics," *The Journal of the Acoustical Society of America*, vol. 133, no. 5, pp. 3072–3082, 2013.

[11] R. Sproat, A. Hunt, M. Ostendorf, P. Taylor, A. Black, K. Lenzo, and M. Edgington, "Sable: A standard for tts markup," in *The Third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis*, 1998.

[12] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: An Update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009. [Online]. Available: http://doi.acm.org/10.1145/1656274.1656278

[13] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.