IOSUD - Universitatea Politehnica Timişoara
Şcoala Doctorală de Studii Inginereşti

CONTRIBUTIONS TO THE SCREENING OF DYSLALIA IN EARLY-SCHOOL-AGED CHILDREN (6-10-YEAR OLDS)

Doctoral Dissertation – Summary

submitted for the Degree of Doctor of Philosophy (PhD)
in the field of Computers and Information Technology at the
Politehnica University of Timisoara

author philol. Emilian-Erman Mahmut
Scientific Supervisor Prof.univ.dr.ing. Vasile STOICU-TIVADAR

July 2024

The doctoral dissertation, titled “Contributions to the screening of dyslalia in the early-school-aged children (6-10-year-olds), consists of 114 pages divided into 6 chapters, incorporating 19 tables, 65 figures, and 77 bibliographic references in support of the research. The main objective of the dissertation is the development of an automated screening algorithm for dyslalic disorders focusing on the following key aspects:

Language Independence: The algorithm should function on the analysis of numeric values extracted from the audio signal, making it applicable across diverse languages. This would eliminate the need for language-specific training data.

Enhanced Efficiency for Speech Language Therapists (SLTs): To support the invaluable work of Speech and Language Therapists, the screening solution focuses on these crucial functionalities:

Rigorous and Consistent Segmentation: The algorithm needs to reliably segment audio samples at the target phoneme level, ensuring consistent analysis across different pronunciations.

Accurate Classification: The algorithm must accurately classify speech samples to identify potential dyslalic disorders.

Accessibility, Cost- and Time-Effectiveness: The solution should leverage open-source resources and be readily accessible in remote areas with limited access to SLTs. This ensures affordability and broad applicability.

The secondary objective is the development of a research infrastructure, which will serve two purposes:

Data Collection and Analysis:
- Anonymized data related to dyslalic disorders in early-school-aged subjects will be collected and stored.
- This data can be utilized to generate valuable reports based on various criteria, including gender, geographical location, and success rates of therapy methods. This information can be extracted from a relational database.

Future research platform: The research infrastructure can serve as a foundation for future research endeavors focusing on speech sound disorders. Researchers can leverage this platform to access data and collaborate on advancements in dyslalic disorder assessment and treatment.

Overall, this research seeks to create a language-independent, accessible, efficient automated screening tool to support SLTs in detecting potential dyslalic disorders. While the screening tool cannot definitively diagnose dyslalia, it can serve as a valuable resource for SLTs in making informed referral decisions and maximizing the impact of their interventions by:

- Improving Detection Efficiency: Automated screening can expedite the process of identifying children at risk for dyslalia.
- Facilitating Early Intervention: Early detection allows for timely referral and intervention, potentially leading to improved outcomes for children with dyslalic disorders

The following section provides a concise overview of each chapter, emphasizing their key contributions to the overall research.

Chapter 1: Setting the Stage - This chapter establishes the foundation for our quest. It introduces the challenges of speech sound disorder (SSD) assessment, explores the potential of automated screening tools, and outlines the specific research objectives that will guide our investigation. Here are the key takeaways:

- Importance of Speech Therapy and Limitations of Current Methods: Traditional speech therapy involves manual assessment, which can be time-consuming and subjective. There is a need for more efficient and objective tools.
- Focus on Dyslalia: The research focuses on dyslalia because it primarily affects the physical production of speech sounds and can be analyzed mathematically.
- The Role of Cognition and Language Acquisition: Speech and language development are interconnected. Brain development and cognitive abilities influence how children acquire and use language.
- Challenges of Speech Segmentation: Speech is a continuous stream of sounds and identifying individual phonemes (speech sounds) can be difficult due to phenomena like phonological assimilation (sounds influencing each other).
- Research Objectives:
    - Develop a language-independent automated screening tool for dyslalia using numeric analysis of audio signals.
    - This tool should efficiently segment phonemes and accurately classify speech samples to identify potential dyslalia.
    - The solution should be accessible, leveraging open-source resources for affordability and wider application.
    - As a secondary objective, a research infrastructure will be built for data collection, analysis, and future research on speech disorders.

The chapter also provides a roadmap for the dissertation, outlining the content of each chapter and how they contribute to the overall research goal.

Chapter 2: A foundation built on knowledge and analysis - Another step in our quest for a sound rationale, this chapter delves into the existing literature on SSD assessment and automated screening methods. It identifies key considerations, limitations in current approaches, and areas ripe for further exploration. This critical review provides a strong foundation for our own research endeavors.

Speech sound disorders (SSDs) are a common challenge for young children, affecting pronunciation and potentially leading to social and academic difficulties. Early intervention is crucial, and computer-based speech therapy (CBST) offers promising tools for speech-language therapists (SLTs).

While traditional methods rely on paper-based assessments, CBST brings advantages like immediate feedback, gamified/storified activities, and data collection for progress tracking. However, concerns exist about screen time, accessibility, and overreliance on technology, highlighting the importance of using CBST as a supplement to traditional therapy.

One key area of CBST development involves automated screening for SSDs. Researchers are exploring how to quantify acceptable pronunciation. This includes mathematically defining the acoustic accuracy needed for clear communication and acknowledging the range of natural variations that do not hinder understanding. Shannon's Communication Theory helps bridge the gap between objective sound analysis and subjective human perception during speech.

Current CBST applications primarily focus on exercises for articulation disorders. However, research is ongoing to address limitations in existing approaches. These limitations include a lack of standardized terminology, the complex nature of assessing SSDs (the need for interdisciplinary approaches), and language-specific variations in pronunciation. Additionally, researchers are exploring how to integrate hearing ability and phonemic awareness (distinguishing sounds) into CBST systems, alongside analyzing real-world sound filtering effects on the way from the speaker's phonoarticulatory apparatus to the listener's auditory cortex.

Recent research has looked at information entropy as a way to address the inherent subjectivity in phoneme articulation. Furthermore, researchers are developing user-friendly ways to present screening results, using visual representations and gamification techniques to reduce user stress and encourage natural speech samples during assessments.

The future of CBST holds exciting possibilities. Researchers are exploring connections between communication quality and dyslalia (speech sound disorders), aiming to leverage data from CBST screenings to understand population trends. Importantly, these tools are envisioned as aids for SLTs, offering advantages like efficient screening, reduced errors, and improved accessibility in areas with SLT shortages. By addressing current limitations and exploring new avenues, CBST has the potential to become a valuable asset in the fight against SSDs.

Our research aligns with existing literature in the following aspects of feature extraction and algorithms:
• Amplitude as a Mispronunciation Discriminatory Feature: supported by [1][2].
• Speech Therapist Utterance as Reference: the speech therapist's pronunciation serves as the reference for comparison with the analyzed phoneme segment [3].
• /r/ Phoneme as Initial Validation Target: The Romanian /r/ phoneme was chosen for initial validation due to its high intensity oscillation characteristics [1][2].

• Targeted Age Range: aligned with the target age group in [3].

To bridge the subjectivity-objectivity gap and boost user motivation our research translates numerical screening results into color-coded isometric visualizations.

The following chapter describes a Feature Extraction and Classification System for audio recordings, developed in accordance with subsets of the research objectives to serve as research infrastructure.

Chapter 3: Building the Tools for Feature Extraction and Classification - Here, we embark on the practical phase of our quest. This chapter details the development of a software application for feature extraction and classification from audio recordings. We built a custom application using C# within the .NET framework to process audio data and validate our research objectives. The exploration of various techniques paved the way for the selection of the most robust and effective methods for our automated screening tool.

Our initial approach focused on amplitude as an indicator of mispronunciation. We compared recordings to those of a speech therapist as a reference point. While we employed polynomial regression for feature extraction and information entropy for classification, this method faced limitations in accuracy and robustness.

To improve accuracy, we explored different polynomial orders for fitting trendlines, increased the detail captured in the features' string representation, and incorporated amplitude information for a richer feature set. We also adjusted the sampling factor to control the level of detail extracted from the audio signal.

We further refined the classification process by adding R-squared calculations to assess how well the model fits the segmented audio portions. Additionally, we implemented automated data recording to expedite processing. While incorporated the logarithmic transformation and Levenshtein distance calculations, these methods were ultimately unsuccessful.

The research prioritized exploring and maintaining flexibility in feature extraction and classification techniques. This application served as the foundation for a more comprehensive speech sound screening system. The iterative development process allowed for continuous improvement of feature extraction and classification techniques.

To ensure cost-efficiency, data security, and scalability, we chose a relational database design. Encrypted storage safeguards personal information. The modular structure allows for seamless integration of future algorithms and functionalities.

The proposed architecture offers a centralized platform for managing speech-language therapists (SLTs), subjects, audio samples, and analysis algorithms. Cloud storage would enhance data accessibility and scalability. The modular design allows for future expansion with new algorithms and interfaces. Potential benefits include improved data management, accessibility, user-friendliness, and cost/time efficiency.

The user interface (UI) caters to different user roles (parents, SLTs, researchers) through role-based access control. The main window acts as a central hub with functionalities tailored to each user type. User-friendly design ensures ease of use for users with varying technical expertise.

The development of the .NET application and its database followed a practical approach, prioritizing functionality over aesthetics to meet the immediate research goals. While the resulting application may appear basic in terms of user interface design and general usability, it successfully delivered the essential functionalities required for validating and invalidating research hypotheses. In essence, chapter 3 details the development of a research infrastructure that serves as the building block for a more advanced speech sound screening system.

The next chapter explores the ongoing quest for a reliable and unified approach to segmenting speech sounds into relevant units.

Chapter 4: Refining the Approach to Segmentation and Classification - An important cornerstone in our quest for a sound rationale, this chapter focuses on refining the segmentation and classification algorithms. It explores various approaches, analyzes their effectiveness, and highlights the limitations encountered. This iterative process allows us to hone our methods and move closer to a reliable automated screening tool.

This chapter explores an entropy-based algorithm for automated screening of /r/ pronunciation in Romanian-speaking children (ages 5-7). The analysis involved audio samples of children's speech productions ("RAFT," "PARĂ," "FAR") collected during speech therapy sessions. An SLT evaluated each pronunciation for accuracy.

The algorithm analyzed the positive amplitude values of the audio signal to generate a polynomial trendline. This trendline was then characterized by assigning "A" (ascending), "D" (descending), or "S" (stable) to each peak and trough. Transition matrices and information entropy values were computed to assess the similarity between the reference (SLT's pronunciation) and test (child's pronunciation) signals. A strict similarity threshold was employed for entropy value difference to identify potential pronunciation errors.

The analysis achieved high match rates with SLT evaluations for initial /r/ (93.3%), with optimal performance at a polynomial order of 11 and S-letter range of 10. Medial and final /r/ showed lower but significant accuracy (80.0% and 83.3% respectively) with distinct optimal configurations. Notably, there was no single parameter setting that consistently yielded the best results across all pronunciations.

The initial approach revealed limitations. While achieving high accuracy in specific configurations, inconsistencies arose across different parameter settings. Efforts to address these inconsistencies through additional application parameters proved unsuccessful. Polynomial orders above 11th and larger S-letter ranges resulted in poorer performance. Feature extraction techniques using logarithmic functions or including maximum/minimum amplitude values did not significantly improve the results. Classification methods based on R-squared values and Levenshtein distance were also ineffective. Chapter 3 provides a detailed description of the enhancements brought to the feature extraction stage. Due to these limitations, the research focus shifted towards developing a more robust algorithm using machine learning for classification.

The following chapter describes a new approach based on cross-correlation for the segmentation stage, which aligns with the established research objectives. Cross-correlation effectively addressed both the phonological assimilation impact and the issue of variable duration of the same utterance by different speakers This approach yielded promising, consistent results that were further refined through the implementation of machine learning classifiers.

Chapter 5: Toward a Promising Solution - Cross-Correlation and Machine Learning - This chapter presents a significant advancement in our quest. It explores a novel two-stage approach utilizing cross-correlation for segmentation and machine learning for classification. The results demonstrate promising accuracy and efficiency, bringing us closer to achieving a robust and reliable automated screening tool.

Stage 1: Segmentation via Cross-Correlation
The system employs cross-correlation, a mathematical technique, to segment individual phonemes from speech recordings. This method effectively handles variations in pronunciation length and complex interactions between neighboring sounds (phonological assimilation).

Unlike traditional methods, cross-correlation does not require manual processing. It automatically segments the targeted phoneme from entire words and generates relevant features for the next stage. Python scripts were written for the segmentation of word-initial, word-medial, and word-final target phonemes. Cross-correlation segmentation offers several advantages:

- addresses the complexities of natural speech, including variable pronunciation lengths and intricate phoneme interactions.
- provides consistent segmentation and generates valuable features for classification.
- automates the process, eliminating the need for manual intervention.

The key findings on Cross-Correlation segmentation (subchapters 5.1-5.3) are the following:
• Cross-correlation effectively segmented the /r/ phoneme in various contexts (vowels, diphthongs, consonants);
• The analysis revealed the influence of phonological assimilation on phoneme characteristics, highlighting the importance of considering neighboring sounds (Subchapter 5.1);
• The segmentation stage proved robust in handling audio samples from different Romanian words (subchapter 5.2).

Stage 2: Classification with Machine Learning
Machine learning models were employed to classify the segmented phonemes as correct or mispronounced. We implemented a selection of machine learning algorithms in Python, including Linear SVM [4][5], Decision Tree (CART)[6][7], AdaBoost [8][9], Gaussian Naive Bayes [10] [11], Multi-Layer Perceptron (MLP) [12], and Logistic Regression [13][14]. This study focused mainly on two models: Linear SVM and Decision Tree (CART). Both models achieved high accuracy (over 97%) in classifying the /r/ sound in various Romanian contexts (words with vowels, diphthongs, and consonants). Notably, the Decision Tree exhibited perfect accuracy on all metrics, while the Linear SVM showed a slight decrease in accuracy with a larger test set. This suggests the Linear SVM might be more sensitive to the amount of training data it receives.

The key findings on Machine Learning classification (subchapters 5.4-5.5) are the following:
• Both Linear Support Vector Machine (SVM) and Classification and Regression Tree (CART) models achieved promising results, with CART achieving perfect scores on some metrics (subchapter 5.5).

• The accuracy of the Linear SVM model improved slightly with a larger training dataset, suggesting potential benefits from increased data (subchapter 5.5).
• While CART showed consistently high performance, the Linear SVM exhibited sensitivity to the size and distribution of training data, highlighting the need for further optimization (subchapter 5.5).

Machine learning classification helped mitigate the challenge of comparing adult female voices (most SLTs are adult women) with children's voices. This is because classification algorithms primarily focus on the spectral envelope and formant frequencies (timbre) of the speech signal, which are less influenced by vocal fold anatomy and pitch differences between adults and children. Training epochs leverage the subjects' data to account for these variations and improve the overall accuracy of the system.

Future work will focus on expanding the dataset size and diversity to improve model performance on unseen data, enabling the system to classify a wider range of phonemes beyond the /r/ sound, integrating language-specific rules to offer more detailed information about mispronunciations, aiding in targeted interventions, and validating the system's effectiveness in real-world clinical settings.

From the overall significance standpoint, this two-stage approach demonstrates potential for developing an automated SSD screening tool. The segmentation stage effectively isolates phonemes for further analysis, while the machine learning models provide promising classification capabilities.

Chapter 6: General Conclusions - The final chapter of our quest reflects on the journey as a whole. It summarizes the key findings, highlights the contributions of the research to the field of SSD assessment, and identifies promising avenues for future investigation.

The dissertation describes a promising system to automatically detect speech sound disorders (SSDs) in children. This two-stage system aims to support Speech and Language Therapists (SLTs) by streamlining the screening process.

The first stage tackles phoneme segmentation, the process of isolating individual sounds from speech recordings. We explored various techniques but found that cross-correlation, a mathematical approach, excelled at segmenting phonemes, particularly the /r/ sound in Romanian. This method effectively factors in variations in pronunciation length and complex interactions between neighboring sounds. Unlike traditional methods, cross-correlation automates segmentation, eliminating the need for manual processing by therapists.

The second stage involves classification. Here, machine learning models take over, analyzing the segmented phonemes to identify potential mispronunciations. The dissertation mainly focused on two models: Linear SVM and Decision Tree. Both models achieved impressive accuracy in classifying correct and incorrect /r/ pronunciations in various Romanian contexts.

This research contributes to the field of SSD assessment in several ways:
• Novel Two-Stage Approach: It proposes a novel two-stage approach for automated SSD screening that leverages cross-correlation for segmentation and machine learning for classification (Chapter 5). This approach offers advantages like consistency, speed, and the potential for language independence.
• Rigorous and Consistent Segmentation: An important contribution of this research is the

development of a rigorous and consistent segmentation stage using cross-correlation (Chapter 5). This approach effectively addresses the following challenges:

- Progressive and Regressive Phonological Assimilation: Cross-correlation can account for the influence of neighboring sounds on phoneme characteristics, including both progressive and regressive assimilation (Chapter 5). This is fundamental for accurate segmentation, as adjacent phonemes influence each other's pronunciation.
- Variable Duration: The cross-correlation method is robust in handling audio samples from different Romanian words with variable durations (Chapter 5). This overcomes the limitation of traditional segmentation methods that struggle with words of differing lengths.

• Machine Learning for Timbre-Related Challenges: The research demonstrates the effectiveness of machine learning in mitigating challenges associated with timbre variations between adult and child voices (Chapter 5). This finding is valuable for developing screening tools that can accurately assess children's speech despite these natural variations.
• Open-Source and Modular Design: The emphasis on open-source resources and a modular design for the software application and database promotes wider accessibility, future expandability, and potential cost-effectiveness for SLTs and researchers (Chapter 3).

This research holds significant promise for improving how SSDs are detected in children. The system offers several advantages:

- Efficiency for SLTs: The automated process translates to faster screening times and potentially reduced workload for therapists.
- Accessibility and Cost-Effectiveness: The use of open-source resources and a modular design promotes wider access and future expansion with minimal additional costs.
- Language Independence: The focus on language-independent feature extraction techniques paves the way for a standardized tool.

The current system primarily focuses on the /r/ sound. Expanding the system's capabilities to classify a wider range of phonemes will enhance its overall usefulness. Additionally, the training dataset used for the machine learning models needs to be increased in size and diversity to ensure the tool performs well on unseen data. Real-world studies are necessary to validate the system's effectiveness in identifying SSDs within clinical settings.

Overall, this research demonstrates the potential of a two-stage approach using machine learning for automated SSD screening.

This dissertation significantly contributes to existing knowledge, evidenced by the publication of 11 scientific papers throughout its development. Notably, six papers are indexed in the prestigious ISI/WOS database. An additional paper is indexed in a valuable Bibliographic Database (BDI) journal, and four other BDI-indexed papers are on track for ISI/WOS inclusion. Furthermore, all publications credit the researcher as the lead author.

Here are the scientific papers that emerged from this research:

1. E. E. Mahmut and V. Stoicu-Tivadar, "Current Challenges in the Computer-Based Assessment of Speech Sound Disorders," in 2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania, May 17-19, 2018, pp. 431-435, doi: 10.1109/SACI.2018.8440970, WOS:000625278800072.
2. E. E. Mahmut and V. Stoicu-Tivadar, "A Speech Sound Disorder Screening System

Database Structure," in Decision Support Systems and Education: Help and Support in Healthcare, J. Mantas, Z. Sonicki, M. Crisan Vida, K. Fister, M. Hagglund, A. Kolokathi, and M. Hercigonja Szekeres, Eds., ser. Studies in Health Technology and Informatics, vol. 255, Zagreb, Croatia, Oct. 15-16, 2018, pp. 185-189, European Federat Med Informat, 2018, doi: 10.3233/978-1-61499-921-8-185, WOS:000455957400036.
3. E. E. Mahmut, M. Della Ventura, D. Berian, and V. Stoicu-Tivadar, "Entropy-based Dyslalia Screening," in Health Informatics Vision: From Data via Information to Knowledge, J. Mantas, A. Hasman, P. Gallos, A. Kolokathi, M. S. Househ, and J. Liaskos, Eds., ser. Studies in Health Technology and Informatics, vol. 262, Athens, Greece, Jul. 5-7, 2019, pp. 252-255, doi: 10.3233/SHTI190066, WOS:000560388600065.
4. E. E. Mahmut, D. Berian, M. Della Ventura, and V. Stoicu-Tivadar, "Optimization of Entropy-Based Automated Dyslalia Screening Algorithm," in Digital Personalized Health and Medicine, L. B. Pape-Haugaard, C. Lovis, I. C. Madsen, P. Weber, P. H. Nielsen, and P. Scott, Eds., ser. Studies in Health Technology and Informatics, vol. 270, Geneva, Switzerland, Apr. 2020, pp. 357-361, European Federat Med Informat, 2020, doi: 10.3233/SHTI200182, WOS:000448144200075.
5. E. E. Mahmut, S. Nicola, and V. Stoicu-Tivadar, "Cross-Correlation Based Automated Segmentation of Audio Samples," in Importance of Health Informatics in Public Health During a Pandemic, J. Mantas, A. Hasman, M. S. Househ, P. Gallos, and E. Zoulias, Eds., ser. Studies in Health Technology and Informatics, vol. 272, Jul. 3-5, 2020, pp. 241-244, doi: 10.3233/SHTI200539, WOS:000630065600062.
6. E. E. Mahmut, S. Nicola, and V. Stoicu-Tivadar, "Cross-Correlation Based Automatic Segmentation of Medial Phonemes," in 2020 14th International Symposium on Electronics and Telecommunications (ISETC), Timisoara, Romania, Nov. 5-6, 2020, pp. 293-296, doi: 10.1109/isetc50328.2020.9301048, WOS:000612681000070.
7. E. E. MAHMUT, M. DELLA VENTURA, and V. STOICU-TIVADAR, “An Entropy-Based Computer Model for the Measurement of Phonetic Similarity: Dyslalia Screening in Early School-Age Children”, Appl Med Inform, vol. 40, no. 1-2, pp. 15–23, Jun. 2018.
8. E. E. Mahmut, S. Nicola, and V. Stoicu-Tivadar, "Decision Tree Versus Linear Support Vector Machine Classifier in the Screening of Medial Speech Sounds: A Quest for a Sound Rationale," Studies in Health Technology and Informatics, vol. 309, pp. 73-77, Oct. 20, 2023, doi: 10.3233/SHTI230742.
9. E. E. Mahmut, S. Nicola, and V. Stoicu-Tivadar, "Support-Vector Machine-Based Classifier of Cross-Correlated Phoneme Segments for Speech Sound Disorder Screening," *Studies in Health Technology and Informatics*, vol. 294, pp. 455-459, May 25, 2022, doi: 10.3233/SHTI220500.
10. E. E. Mahmut, S. Nicola, and V. Stoicu-Tivadar, "Word-Final Phoneme Segmentation Using Cross-Correlation," *Studies in Health Technology and Informatics*, vol. 275, pp. 132-136, Nov. 23, 2020, doi: 10.3233/SHTI200709.
11. E. E. Mahmut, S. Nicola, and V. Stoicu-Tivadar, "A Computer-Based Speech Sound Disorder Screening System Architecture," *Studies in Health Technology and Informatics*, vol. 251, pp. 39-42, 2018, doi: 10.3233/978-1-61499-921-8-39.

References

[1] Grigore, O., Grigore, C., Velican, V., Intelligent System for Impaired Speech Evaluation, Recent Advances in Circuits, Systems and Signals, Book Series: International Conference on Circuits Systems Signals, ISBN:978-960- 474-226-4, WOS:000290360900078, Malta, Sept. 2010, p. 365-368.
[2] Grigore, O., Velican, V., Self-Organizing Maps For Identifying Impaired Speech, Politehnica University of Bucharest, 061071, Romania, Department of Applied Electronics and Telecommunications Faculty, in Advances in Electrical and Computer Engineering (AECE), Vol. 11, No. 3, 2011;
[3] Pentiuc, Ș. G., Schipor, O. A., Gîză Belciug, F., Belciug, C. E., Nestor, M., Aplicație Logomon - Raport final Proiect TERAPERS, Sistem pentru terapia personalizată a tulburărilor de expresie lingvistică. Suceava: Universitatea Ștefan cel Mare, Facultatea de Inginerie Electrică și Știința Calculatoarelor, 2006;
[4] Mahmut, E.E., Nicola, S., Stoicu-Tivadar, V., Support-Vector Machine-Based Classifier of Cross-Correlated Phoneme Segments for Speech Sound Disorder Screening, Stud Health Technol Inform., 294, p. 455-459, May 2022;
[5] Kecman V., Learning and Soft Computing, Support Vector Machines, Neural Networks, and Fuzzy Logic Models, Cambridge, Massachusetts, The MIT PRESS, p. 148-166, 2001;
[6] Kumar, P., Kumar, D., Decision tree classifier: a detailed survey, International Journal of Information and Decision Sciences, Vol. 12, No. 3, p. 246-269, Apr. 2020;
[7] Mahmut, E.E., Nicola, S., Stoicu-Tivadar, V., Decision Tree versus Linear Support Vector Machine Classifier in the screening of medial speech sounds: a quest for a sound rationale, Stud Health Technol Inform., 309, p. 73-77, Oct. 2023;
[8] Ding, Y., Zhu, H., Chen, R. and Li, R., An Efficient AdaBoost Algorithm with the Multiple Thresholds Classification, Appl. Sci., 12, 5872, Jun. 2022;
[9] Wang, W. and Sun, D., The improved AdaBoost algorithms for imbalanced data classification, Information Sciences, Volume 563, p. 358-374, Jul. 2021;
[10] Tieppo, E., Nievola, J.C. and Barddal, J.P., Adaptive learning on hierarchical data streams using window-weighted Gaussian probabilities, Applied Soft Computing, Vol. 152, Feb. 2024;
[11] Chen, S., Webb, G.I., Liu, L. and Ma, X., A novel selective naïve Bayes algorithm, Knowledge-Based Systems, Vol. 192, March 2020;
[12] Chatterjee, A., Saha, J. and Mukherjee, J., Clustering with multi-layered perceptron, Pattern Recognition Letters, Vol. 155, p. 92-99, Mar. 2022;
[13] Hu, W., Qian, Y., Soong, F.K. and Wang, Y., Improved mispronunciation detection with deep neural network trained acoustic models and transfer learning based logistic regression classifiers, Speech Communication, Vol. 67, p. 154-166, Mar. 2015;
[14] Siniscalchi, M.S., Combining speech attribute detection and penalized logistic regression for phoneme recognition, Neurocomputing, Vol. 93, p. 10-18, Sept. 2012;