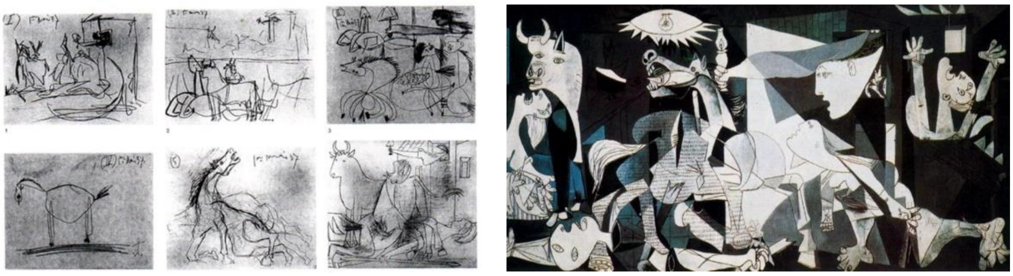
Message from editor

Dear respected authors and readers,

It is the 7th number of our “Works in Progress in Embedded Computing Journal (WiPEC)”. The Journal is dedicated to the Special Session “Works in Progress” within MECO and CPSIoT conferences. Topics of interest of the WiPEC include, but are not limited, to the topics covered by MECO and CPSIoT, see at www.embeddedcomputing.me .

There is a specific beauty in sharing work in progress and there is a risk in seeking finished perfection in all that we as scientists do. Picasso had a definite view on finishing his artworks 'Woe to you the day it is said that you are finished! To finish a work. To finish a picture. What nonsense! To finish it means to be through with it, to kill it, to rid it of its soul – to give it its final blow; the most unfortunate one for the painter as well as for the picture.’ He also identified with the role of accidents not as mistakes but as an inevitable part of the process and the path to discovering humanity in art 'Accidents, try to change them - it's impossible. The accidental reveals man”.



The Work in progress - Study for "Guernica" and the Completed Guernica

As a scientist I do believe that science is a never-ending process because we can never achieve complete certainty in our understanding of how the universe works. The scientific process is never done, mainly because we can never rule out the possibility that even our most well-supported models may not hold under every set of circumstances. And the beauty of this incompleteness sparks further curiosity and investigations. The whole scientific journey which we are embarking on is a work in progress. As an educator I do believe that we increasingly deal with “known unknowns”. These are things we know will change in the future, but their exact form is unknown. Artificial intelligence, a subject often researched in this WiP, will undoubtedly play an important role.

The COVID-19 pandemic has generated an extraordinary amount of data in a short time period, enabling the rapid generation and dissemination of related research. More than a year after the declaration of a worldwide pandemic, the origin of Covid-19 is still a “work in progress”.

The “Works in Progress” Journal dedicated to the Special Session “Works in Progress” within MECO and CPSIoT conferences celebrates the steps taken towards learning as much as it celebrates the finished product. MECO2021 values intellectual the curiosity and the processes of learning, thinking both critically and innovatively that lead to the final product. Being reflective, honest with one’s own research and its limits and perspectives, and sharing knowledge are exactly the values supported and celebrated by the WiP Journal. It celebrates the small steps of man and giant leaps of mankind in terms of scientific achievements and intellectual risks into the scientific unknown.

Here we celebrate the love that the authors have for learning, the many stepping-stones that have been crossed to gain success. We celebrate that knowledge, much like the universe, is infinitely expanding and that knowledge expansion is only completely successful if we search for explanations of concepts and ideas which aren't completely understood- all of which is celebrated through the values of the MECO and ECYPS conferences.

Thanks to all of the contributors!

Sincerely yours,

Editor

Prof. Gordana Laštovička – Medin, PhD

University of Montenegro

Utilization of Low-Cost Sound Sensors with a built in Microphone as a Respiratory Pattern Sound Indicator and a Risk Mitigation Tool

In response to COVID-19

Gordana Laštovička-Medin
University of Montenegro
Podgorica, Montenegro
gordana.medin@gmail.com

Rajka Pejanović
University of Montenegro
Podgorica, Montenegro
p.r.rajka@gmail.com

Abstract—**Since the detection of pattern abnormalities may lead to not only the prevention of chronic respiratory diseases but also other diseases, many techniques have been developed in order to detect breathing and coughing patterns. To benefit from the cross-disciplinary studies we have decided to expose physics students to both: learning about sound using coughing as a targeted research topic and to develop a demo tool that is useful for building on exploratory skills and provides them with solid knowledge for future more advanced scientific research in biomedical engineering. A low-cost microphone sensor was tested for the purpose of understanding whether it can be used not only as a sound indicator but more broadly as a risk mitigation tool during a pandemic such as the current pandemic, COVID-19. The final goal of this long-term project is to build mathematical models aiding the identification of features from sound samples and to apply a classifier algorithm based on the machine learning technique at the final stage of research.**

Keywords- COVID-19, sound pattern, human generated cough and breathing, low-cost sound sensor, loudness

I. INTRODUCTION

COVID-19 is a respiratory infection caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. Prominent symptoms of COVID-19 include coughing and breathing difficulties. Cough sound analysis helps us to differentiate two similar sounds and to define the objective correlations with spirometry and clinical diagnosis [2], including Cough Peak Flow using cough sounds [3]. The auscultation of the respiratory system is another diagnostic technique and an inexpensive, noninvasive, safe, and easy-to-perform method [4].

The parameters such as frequency, intensity, and timbre of sound are of particular interest for the classification of respiratory diseases and are defined as follows. Pitch is the subjective perception of sound's frequency and depends on the frequency while amplitude of loudness is related to the energy of sound waves and is measured by the height of sound waves from the mean position; it is the subjective perception of amplitude. Quality or timbre is an important property of sound that differentiates two sounds with the same pitch and loudness. The fundamental frequency or primary frequency is the lowest frequency of a sound wave and it determines the pitch of the sound; the frequencies higher than the fundamental frequencies are called overtones while harmonics are overtones whose frequencies are whole number multiples of the fundamental frequency.

However, real-world sounds are not usually deterministic: they do not just have simple harmonics of the fundamental frequency. Instead, they also have unpredictable “inharmonic” frequencies that are not structured as noise. Thus having complete understanding of these measurable quantities and designing an experiment where those features are not lost when recorded and processed is crucial for further applications of machine learning techniques [5]. The issue is also how to deal with the research complexity without compromising the flexibility of techniques required for the extraction of sound features and still providing a comprehensive outcome that would not compress important information for the sake of data reduction. Towards this, this paper presents an early effort, mostly exploratory based, in building the capacity for such a complex and comprehensive task and towards creating a cough/respiratory sound database in Montenegro.

II. EXPERIMENTAL PROCEDURE

A. *Research question*

The cough frequency is the most basic measure of coughing, but the objective study of cough signals has the potential to identify further features which may be clinically relevant and hence useful endpoints to study. Here we recall an early measurement (Fig. 1a) that used the audio tapes as research tools where the behavioral changes were monitored in order to extract untypical patterns over a longer time period [6]. Quantifying a cough was never an easy task, it is still not fully understood, and the symptoms are often incorrectly assigned. There is also no universally agreed unit of cough. The most intuitive way to quantify cough is to count the expulsive (first phase) of cough sounds (Fig. 1b). If long bursts (or peals) of expulsive cough sounds are present, then to identify each expulsive phase can be exceedingly difficult. The temporal patterns of coughing vary both in the short term (peals or epochs of coughs versus single coughs) and in the long term,

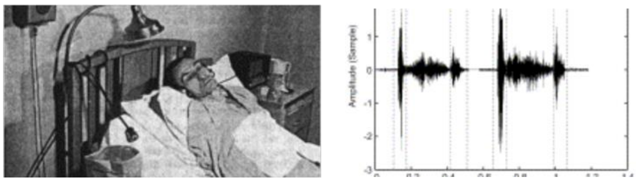


Figure 1. a) Early objective cough monitoring study with wall mounted microphone recording sounds from a hospital in-patient; b) Typical cough waveforms with expulsive phase, intermediate phase and voiced phase.

In another research project the respiratory sounds were recorded simultaneously with 16 microphones distributed over the thoracic surface [7]. Acoustic energy in three frequency bands, 150-300 Hz, 300-600 Hz and 600-1200 Hz, was analyzed during inspiration and expiration and results indicate that on average, inspiratory sounds are 10 to 11 dB louder than expiratory sounds at comparably flow rates. Findings that the acoustic shadow produced by the heart is more pronounced during inspiration support the concept that inspiratory sounds are produced predominantly in the periphery of the lung, while expiratory sounds are generated more centrally; lung sounds were also found to be significantly louder on the right side for the front part of the thorax, and on the left side for the back suggesting the importance of sensor positioning on the accuracy of the diagnostic procedure. [7]. Auscultation of the infant's chest reveals that their lung sounds seem to be different to those of adults since the normal lung sounds of newborn infants contain higher-frequency components than those of adults as a result of less filtering of the lung sound in infants [8]. Furthermore, accompanying airflow limitation or poor transmission of sounds was found for pulmonary emphysematous, indicating role of sound intensity as an important feature to be understood. From the point of view of the mechanism the sound is generated and then blocked over the transmission journey [9].

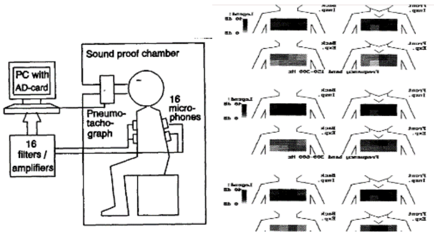


Figure 2. a) Sound loudness explored via microphones attached a 16 chest; b) The upper 4 diagrams represent the low frequency band, the middle 4 diagrams the intermediate frequencies and the lower 4 diagrams the high frequency band position. Microphones on the thoracic surface were arranged in a strictly geometric pattern to allow a meaningful comparison between different parts of the thorax [7].

It is important to also be aware that the testing of any new cough monitoring systems needs to be rigorous. Unfortunately, even the most recent medical systems suffer from a lack of sufficient accuracy while inconsistency in performance adds additional uncertainty in sound interpretation. It is also well known that any cough monitor will identify some cough events correctly (*true positives*), mistake non-cough events as coughs (*false positives*), miss some cough events (*false negatives*) and correctly ignore non-cough (*true negatives*).

All those problems and uncertainties in designing the data taking approach motivated us to set up our own investigation seeking for answers in long and short term. The research question is not only to experimentally test and demonstrate whether we can use a low-cost sound sensor module - MAX4466 Microphone Amplifier Module with an adjustable Gain Breakout Board for Arduino to detect sound patterns and utilize this tool for risk mitigation during COVID-19, or any respiratory related diseases, but also to learn more about the way the errors are overlooked and ignored (such as differentiation from ambient noise, differentiation from other patient sounds, especially speech, laughing sneezing and variability in the acoustics of cough sounds between individuals including transmission path of sound, positioning of producer of sound etc.).

B. Waveform patterns

Here we present some known waveform patterns that we will use to mimic "known/classified" sounds when requesting healthy volunteers to imitate the sound pattern. The different duration and pattern behavior of the three acoustic phases called the explosive, intermediate, and voiced (right) phases (introduced in the previous section) are clearly observed. Vesicular breath sounds (Fig. 3a) are heard across the lung surface. They are lower pitched, rustling sounds with higher intensity during inspiration. During expiration, sound intensity can quickly fade. Inspiration is normally 2-3 times the length of expiration [10],[12]. Bronchial breath sounds are tubular, hollow sounds which are heard when auscultating over the large airways (e.g. second and third intercostal spaces). They will be louder and higher pitched than vesicular breath sounds (Figure 3b). Figure 4a shows the wave form produced when one wheezes. As seen, wheezes can be high or low pitched. The waveform for crackles is displayed in Figure 4b. The pitch is lower than late inspiratory crackles. A patient's cough may decrease or clear these lung sounds.

Figure 3. Wave forms for: a) Vascular and b) bronchial breath.[10]

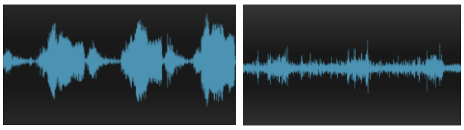


Figure 4. Waveform plot: a) Wheezes. b) Crackles [10]

C. Set up

The research tool was simple: it was based on Arduino and a microphone sensor: MAX4466 Microphone Amplifier Module, with electret microphones on it as shown in Figure 5. This breakout is applicable in projects such as voice changers, audio recording/sampling, and audio-reactive projects that use FFT. The Electret Mic Breakout translates amplitude (not volume) by capturing sound waves between two conducting plates (one a vibrating diaphragm and the other fixed) in the microphone and converting them into electrical waves. These electrical signals are then amplified and picked up by the microcontroller's ADC

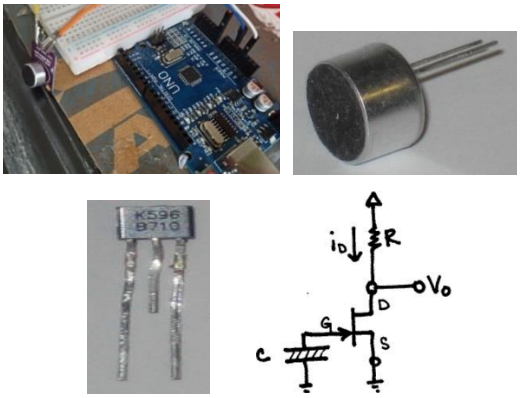


Figure 5. a) Sound sensor with microphone built in connected to the Arduino; b) Electret microphones c) Amplifier transistor (2SK596); d) Equivalent schematic of electret microphone.

Electret microphones: The amplifier consists of a single JFET transistor, with the gate connected to the pick-up plate, the source connected to ground, and the signal appearing on the drain. This is called a common-source configuration, as the source is connected to ground, which is common to all signals. The JFET in this electret microphone is a 2SK596, which is designed for low-noise applications. The electret maintains a fixed charge, and therefore maintains a voltage across the capacitor. For details we refer to [11].

Measuring Sound Levels: The Audio signal from the output of the amplifier is a varying voltage. To measure the sound level, we had to take multiple measurements to find the minimum and maximum extents or "peak to peak amplitude" of the signal. We have chosen a sample window of 50 milliseconds. That is sufficient to measure sound levels of frequencies as low as 20 Hz - the lower limit of human hearing. After finding the minimum and maximum samples, we computed the difference and converted it to volts and the output was printed on the serial monitor. We also did some experimental research with sounds at different volume levels to test how our average, min, max and span values respond. Adjustment to the the gain potentiometer was occasionally needed in order to utilize the max span for our sound levels while not overdoing it and not affecting the +/- sign of data.

D. Calibration:

The first step was to calibrate the sensor and to define the noise level (when no sound was produced). This procedure required additional adjustments of electronic components to provide optimal sensor performance. The optimal choice would be to have an acoustically isolated room which we were unable to obtain partially due to limits imposed by COVID-19 measures. Figure 5 shows the recorded noise level.

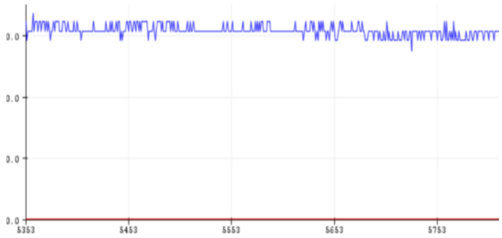


Figure 6. Sensor calibration: Noise level

E. Row Data, Sampling, Response & Limits in sensor performance

It is important to note that sound patterns are generated by volunteers, thus not clinically confirmed, and not taken from any approved medical data. For sampling the rule is to sample twice as fast as the maximum frequency (20 KHz) we want to capture. As we increase our knowledge about microphones and artefacts, we will learn more about the limits of its application in our research. Since, a JFET is used as the amplifier and because it has a high input resistance (30 MΩ or more) this means that almost no current is pulled off the electret capacitor. The reason for this is that if the amplifier had a lower input resistance, the low frequency response of the microphone would suffer. This is because the input stage acts like a high-pass filter, with the electret being the capacitor, and the input of the amplifier being the resistor, and larger values of R and C give lower cut-off frequencies. The main noise sources in this microphone are pick-up noise and transistor noise. Since the entire capsule is sealed and grounded, the pick-up noise is very low and usually not noticeable. The transistor noise, on the other hand, can be quite high, due to the high input resistance on the JFET. Typical values are around -120 dB to -110 dB, which may sound rather low, but the audio signal level is usually less than -40 dB, so it's only an 80 dB signal to noise ratio (SNR). This is a common issue with condenser microphones due to the high input resistances required. Regarding the sensor sensitivity we learnt that a smaller diaphragm would tend to give better high frequency and distortion characteristics, but will not be as loud, and therefore have worse SNR. A smaller diaphragm, as we found later in specifications, will also have a smaller capacitance, so its low frequency response will not be as good. Thus, picking a certain microphone for application must be rigidly tested.

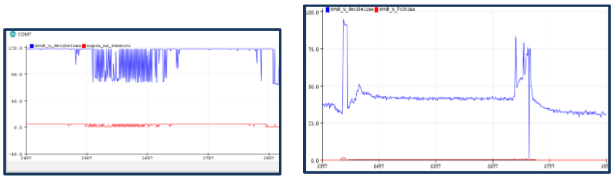


Figure 7. a) Raw data: after a period of silence a burst of cough is produced, and pattern repeated. b) Strong short burst of cough, then repeated after a prolonged time of silence. Some artefact added in the second burst.

Figure 7a show the relative sound level that the microphone picks up. The multiple samples were taken during a sample window which is set to 25 mS (50 mS = 20 Hz). The amplifier in this sound module is biasing the output at 1/2 of the Vcc used to power the board. The drawback here is that any audio (AC voltage) received and amplified will just randomly add or subtract from that 'center' value. Different scenarios for sound generation have been created. Fig. 7a) shows sound level picked up by the microphone after a certain period of silence was prolonged, then burst of coughs was produced, and this pattern further repeated. In Figures 7b) we slightly modified the scenario: a healthy male volunteer, after a period of being silent, (breathing quietly) generated strong but short coughs. We tried to mimic the crackles, but it becomes obvious that supervised training is required in order to obtain an optimal simulated pattern for crackles. However, the sensor was able promptly to record the change in sound pattern.

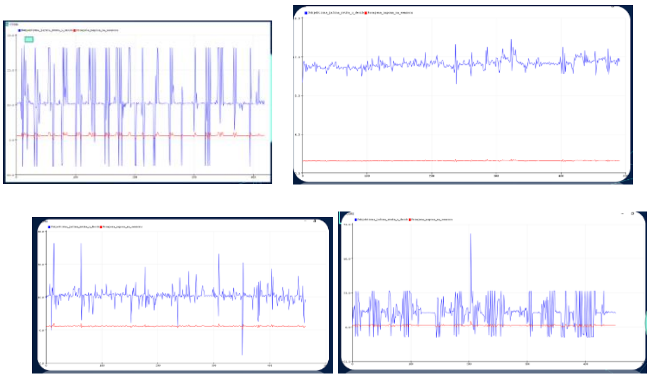


Figure 8. Raw data: a) Upper left image - Cough prolong over longer time, b) upper right image - monotonous breathing and c) bottom image - rapid breathing; d) singing; in the middle of singing a sharp clap was added (seen in image as spike)

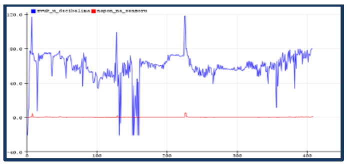


Figure 9. Randomly generated sound: Beathing, Coughing plus some artefacts.

The prolonged cough with the almost uniform intensity, the monotonous breathing, and the rapid breathing, are shown in, Figure 8a, 8b, and 8c, respectively. The variation of loudness while a person was singing is displayed in Fig 8d) while more randomly by human generated sound (in the pattern repetition and the amplitude variation) where breathing, short but strong coughing and for a noticeably short timing interval induced non-vocal artefact-sounds such by clapping, or screech (sharp spike) as presented in Figure 9.

F. Data Processing

The previous section was devoted to Data Acquisition. Approaching the pitch extraction is more challenging than the envelope extraction. The zero crossing technique was prone to errors because of harmonics causing additional zero-crossing. Techniques based on picks of the filtered sample might be a better option, but we struggled with the processed sample (only one we tested) since the fundamental frequency was week. The outcome of an FFT due to large "inharmonics" was difficult to interpret. However, we learnt a lot and we are currently working on gaining a better understanding of the techniques and the sampling. Regarding the envelope extraction technique, performed off-line, some issues had to be solved before we could proceed with the analysis. The cable that was used to connect the sensor to the computer had an issue and we solve it by providing better shielded. The noise that was constantly present in surrounding, was identified, recorded, then inverted in sign and added to the recorded sample. The artefacts affecting the duration of inspiration and expiration phases cause irreproducibility of the envelope extraction, and this is an internal feature of the signal envelope definition. An envelope of a signal is what is obtained through tracking successive ("connected") peak values and the technique can be described as follows: firstly, the signal is squared, then it is passed through a low pass 3rd order filter and in the final step the square root of the signal obtained from the previous step is calculated. Since, the extraction of the envelope is very useful for the classification of data by exploiting the features such as duration of inspiration and expiration phases giving us indications whether the analyzed data correspond to, for instance the crackled breathing where breath sounds are discontinuous and non-musical or to another respiratory disease, the control over all sources of uncertainty is crucial.

Many sound software for laptops can be used free of download but we suggest testing them first. Our first results on envelope extraction used data taken from a volunteer by requesting him to imitate crackled breathing with varying loudness and sampling was perform with frequency of 20 KZ (Figure 10). The expulsive phase and intermediate phase of cough is not distinguished well; from this we learnt that more work on filter algorithms has to be done. Some systematic shift was observed too (Fig 10b).

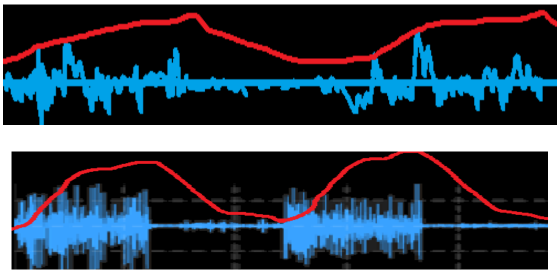


Figure 10. First insight into signal envelope extraction: a) upper image: volunteer mimic quite crackle breathing but sensor was occasionally lost the contact to computer; b) bottom image: Volunteer was asked to produce burst of cough and repat it. The systematic shift can be observed. More work on understanding the algorithm and outcome of analysis is undergoing.

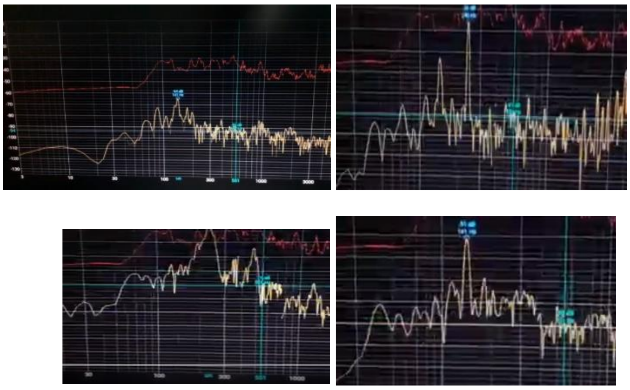


Figure 11. Power spectral density presented by blue line. The figure is the result of Fourier analysis and it represents the distribution of the intensity per frequency. The red line was left from previous measurement. The data filtering was not studied.

Few problems we identified. One issue is that we are converting an analogue input into a digital pin and this means that the triggering is not performed at a consistent voltage except if we connect the audio signal into a voltage comparator before it is converted into a digital signal through a digital pin; this way the pulse is always triggered at the same point or the same threshold voltage. Another thing is that the “envelope” only gives the information about the loudness, so one has to keep in mind that there is no frequency information in the extracted envelope. The real sounds are complex wave shapes, and it is hard to pick out when one cycle ends and the next starts, something that is relatively easy with the mathematically defined wave shapes such as squares. What makes it even harder is the fact that real sounds change their harmonic content and wave shape as the sound continues. One can learn a lot about envelope from analysing music notes played on string instruments and comparing the four envelope segments: the start of the sound is called the “attack” segment; after the loudest part of the sound, the fall to a steady “sustain” segment is called the “decay segment and when the sound ends, the fall from the sustain segment is called the “release” segment. Length of those segments can be used to differentiate different sources of sound; however, using only extracted envelope as feature for a certain type of cough seems to bring a lot of ambiguity. As excesses we play around the vibration, loudness and timbre of sound produced on cello and look at the change in the shape of sound envelope.

The quick FFT as a demo to students was performed without using an advanced data filtering approach, since the aim was to visualise change in frequency response as sound continue. An example of a power spectral density of a cough signal is shown in Fig. 11. The visible spikes present artefact. In future work the signal will be passed through the band pass filter to get rid of the low frequency noise. In this way we will eliminate most of the background noise present in surrounding including electronic sound of laptop ventilation. This will allow us to extract features such as the calculated length of a cough, the length between two hits in a cough attack and the total length of the cough sequence in processed data. It is also important to know that sound heard from mouth and those listened by practitioners from stethoscopes in auscultation differ from each other in the frequency content since some parts of body behave as low filters. This as an interesting topic to be further explored. We also found that some frequencies assigned as classifiers to some respiratory problems has overlapping with some musical notes (depends on what an octave is covered); so, the scope for a false positive is vast. We also noticed that some audio signals might have noise on them or harmonics in the waveform that cross the threshold voltage many times rapidly within the cycle we measured.

Since the MAX4466 is "optimized" for extremely low power, it might not be a good choice for comprehensive sound analysis, but it can be used as an experimental and learning tool. By visualizing what is lost and what information is gained using Fast Furrier Transformation with and without applying the data filtering one could learn more about changes in harmonic content and timbre (and resolution vs sample size which again indicate necessity for the good signal filtering due to huge processing time). On the other hand, the timbre that helps to distinguish two sounds with the same loudness depends on the relative strengths of the components of different frequencies and is mainly determined by the harmonic content of a sound and the dynamic characteristics of the sound such as vibrato and the attack-decay envelope of the sound. Thus, one has to be careful as to what extent the quality of original information or original content may be lost in data processing.

As a future exercise we might decide to use an area of microphone sensors adjusted to a certain frequency band using voltage control. Different microphone sensors will be explored too. Additionally, sensors such as a temperature, pressure sensor and an accelerometer attached to the chest would be an interesting and promising setup as well as an interesting new idea for those looking for more innovative and cross-disciplinary teaching approaches at the University.

III. Conclusion

In our study we explored the utilization of low-cost sound sensors with built-in microphones. We managed to demonstrate that our few-pound sound sensor is sensitive enough to distinguish different sound amplitudes and to precisely record the oscillation in sound pattern behavior.

Originally, we started this project as student lab exercise in response to COVID-19, and to test the proof of concept if a cheap microphone sensor could be useful (and to what accuracy level) in pandemic times (such as COVID-19) as a quick solution for risk mitigation (not applicable for clinical application before rigorous tests are performed). The presented analysis could be promising when it comes to differentiation of various sounds coming from the patient. The results are good starting point for further development of a solution that can be widely used for clinical purposes too.

The presented project at this level of development can be immediately utilized as a useful demo educational tool. The envelope extraction technique is also developed, and the first preliminary results are currently under inspection. More work on FFT as well as understanding the extracted envelope is undergoing. For this our priority is to develop a more comprehensive and robust filtering algorithm. After that we will process data to extract respiratory rate (RR). This would allow us to extract the features from the signal in form of vital signs important for COVID-19.

This research would further benefit from the inclusion of a targeted group of people with chronic respiratory problems. More data is needed to give a conclusion about the reliability of the sensor and its usage as a medical monitor device. However, preliminary tests clearly demonstrate that sensor is capable to follows variation in sound pattern indicating further that such a simple and low-cost set-up would be an efficient tool for sound mapping and sound monitoring of a certain environment and thus can be suitable aid for deaf people also during respiratory pandemics such as COVID.

We are also fond of sharing our experiences with readers. Open questions and issues we found have motivated us to work further on this topic. Filtering of the frequencies and harmonics during sound journey before it is captured by the sensor is also an interesting topic and research on it would give us more knowledge about lost information that might be useful in sound classification and the clarification of the uncertainties in its association to different respiratory diseases or other diseases. The published papers usually ignore the sources of uncertainty and we found it useful to open such questions.

Acknowledgment

We want to thank the MECO organizers for supporting our work by managing the continuation of the conference during the extremely challenging time of COVID-19 pandemic.

References

[1] https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[2] Gowrisree Rudraraju and Sai Praveen Haranathc, “Cough sound analysis and objective correlation with spirometry and clinical diagnosis”, Informatics in Medicine Unlocked, Volume 19, 2020, 100319.

[3] Yasutaka Umayahara et al.,“Estimation of Cough Peak Flow Using Cough Sounds”, Sensors (Basel),. 2018 Jul 22;18(7):2381., doi: 10.3390/s18072381.

[4] Malay Sarkar, Irappa Madabhavi, Narasimhalu Niranjan, and Megha Dogra, “Auscultation of the respiratory system, Ann Thorac Med”. 2015 Jul-Sep; 10(3): 158–168., doi: 10.4103/1817-1737.160831

[5] Hans Pasterkamp, Steve S. Kraman and George R. Wodicka, “Respiratory Sounds Advances Beyond the Stethoscope”, American Journal of Respiratory and Critical Care Medicine, Volume 156, Issue 3, PubMed: 9310022, 1997

[6] Jaclyn Smith, “Ambulatory methods for recording cough”, Pulmonary Pharmacology & Therapeutics, Volume 20, Issue 4, August 2007, Pages 313-318

[7] Martin Kompis and George R. Wodicka, “Distribution of inspiratory and expiratory respiratory sound intensity on the surface of the human thorax”, Proceedings - 19th International Conference - IEEE/EMBS Oct. 30 - Nov. 2, 1997 Chicago, IL. USA

[8] J F Kanga, S S Kraman, “Comparison of the lung sound frequency spectra of infants and adults”, PMID: 3774386 DOI: 10.1002/ppul.1950020508

[9] Schreur HJ et al., “Lung sound intensity in patients with emphysema and in normal subjects at standardized airflows”, DOI: 10.1136/thx.47.9.674 PMID: 1440459 PMCID: PMC474797

[10] https://www.easyauscultation.com/vesicular-breath-sounds

[11] https:/www. openmusiclabs.com/learning/sensors/electret-microphones

[12] https://www.easyauscultation.com

MENOnet JOURNAL: WORKS in PROGRESS in EMBEDDED COMPUTING (WiPiEC), Volume 7, Issue 1, JUNE 2021

Classification of pedagogical content using conventional machine and deep learning model

Vedat Apuk, Krenare Pireva Nuci
Department of Computer Science and Engineering
University for Business and Technology
10000 Prishine, Kosovo
Email: krenare.pireva@ubt-uni.net

Abstract—**The advent of the Internet and a large number of digital technologies has brought with it many different challenges. A large amount of data is found on the web, which in most cases is unstructured and unorganized, and this contributes to the fact that the use and manipulation of this data is quite a difficult process. Due to this fact, the usage of different machine and deep learning techniques for Text Classification has gained its importance, which improved this discipline and made it more interesting for scientists and researchers for further study. This paper aims to classify the pedagogically content using two different models, the K-Nearest Neighbor (KNN) from the conventional models and the Long short-term memory (LSTM) recurrent neural network from the deep learning models. The result indicates that the accuracy of classifying the pedagogical content reaches 92.52 % using KNN model and 87.71 % using LSTM model.**

Keywords: Document Classification, KNN, LSTM, coursera dataset, education, text classification, deep learning models, machine learning models

I. INTRODUCTION

Billions of users create a large amount of data every day which in a sense comes from various types of sources. This data is in most cases unorganized and unclassified and is presented in various formats such as text, video, audio, or images.

Processing and analyzing this data is a major challenge that we face every day. The problem of unstructured and unorganized text dates back to ancient times, but Text Classification as a discipline first appeared in the early 60s, where 30 years later the interest in various spheres for it increased [1], and began to be applied in various types of domains and applications such as for movie review [2], document classification [3], ecommerce [4], social media [5], online courses [6, 7].

As interest has grown more in the upcoming years, the uses start solving the problems with higher accurate results in more flexible ways. Knowledge Engineering (KE) was one of the applications of text classification in the late 80s, where the process took place by manually defining rules based on expert knowledge in terms of categorization of the document for a particular category [1]. After this time, there was a great wave of use of various modern and advanced methods for text classification, which all improved this discipline and made it more interesting for scientists and researchers, more specifically the use of machine learning techniques. These techniques bring a lot of advantages, as they are now in very large numbers, where they provide solutions to almost every problem we may encounter.

The need for education and learning dates back to ancient times, where people are constantly improving and trying to gain as much knowledge as possible. There are various sources of learning available today including various MOOC platforms such as Coursera, Khan Academy, Udemy, Udacity, edX, to name a few, and as technology has evolved it has contributed to better methods of acquiring knowledge that will facilitate this process. The data coming from these sources are in most cases in digital form, more specifically in the form of video and text lessons. The platforms that contain these lessons are called Massive Open Online Courses (MOOCs), where in addition to the video lesson, it also contains its textual representation called a transcript. Considering that the duration of a video lesson depends on several parameters, such as the category of video material, the platform on which the lesson is provided, the complexity of the topic, the number of instructors, and the group of lesson attendants. The duration of the lessons indirectly dictates how long the transcript will be, in other words how many words it can contain. The category shows the nature of the video and the topics that will be presented in it. As it is already known, that each video lesson belongs to a certain category, or in a group of categories, so does the transcript as well. Given this advantage, we can conclude the fact that text classification is becoming quite extensive as a discipline, where also its use can solve many challenging problems in every domain, and specifically in education domain.

The aim of this paper is to investigate two classification techniques that are used to classify the pedagogical content, and the focus tends to compare conventional machine learning models with deep learning models, by selecting KNN algorithm for the first approach and LSTM architecture for the latter one.

To better indicate the idea we want to present, the paper will be divided into several sections, as follows: as part of literature review the main processes of classifying documents are explained, continuing with related work conducted so far in this area. In the experimental section, the design of the conventional machine learning models and deep learning models will be elaborated and the results for the each of the architectures will be presented using a number of evaluation techniques (recall, precision, F-Score, accuracy). The paper will be concluded with conclusion and future work.

II. RELATED WORK

Text mining or text analytics is one of the artificial intelligence techniques that uses Natural Language Processing (NLP) to transform unorganized and unstructured text into an appropriately structured format that will make it easier to process and analyze data. For businesses and other corporations, generating large amounts of data has become a daily routine. Analysis of this data help companies gain smarter and more creative insights regarding their services or products collected from a variety of sources in automated manner. But this analysis step requires processing a huge amount of data where the data needs to be prepared, and this is in most cases the cause of various problems.

NLP consists of five steps or phases (see Figure 1), such as: Lexical Analysis, Syntax Analysis, Semantic Analysis, Pragmatics, and Discourse [8].

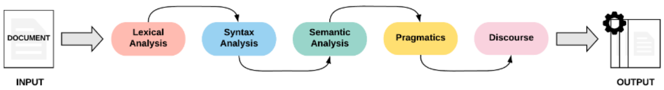


Figure 1: Natural Language Processing steps.

So, the goal of text classification or text analysis is to structure and classify data to facilitate the analysis process. Today, as shown in Figure 2, in order to perform text classification in the existing data, we follow the four phases [9]:

a) *Feature Extraction,*
b) *Dimension Reductions,*
c) *Classifier Selection,*
d) *Evaluation.*

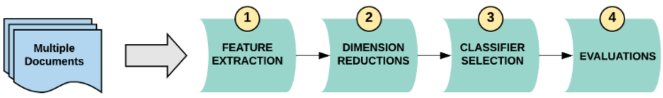


Figure 2: Four-phase model of a text classification system.

As shown in Figure 2, with feature extraction as an initial phase one piece of text or document is converted into a so-called structured feature space, which will be useful to us when using a classifier. But prior to this, needs to perform data cleaning, taking care of missing data, removal of unnecessary characters or letters, in order to bring the data in an appropriate shape for extracting the features, otherwise omitting the data cleaning can directly affect negatively the performance and the accuracy of the final results.

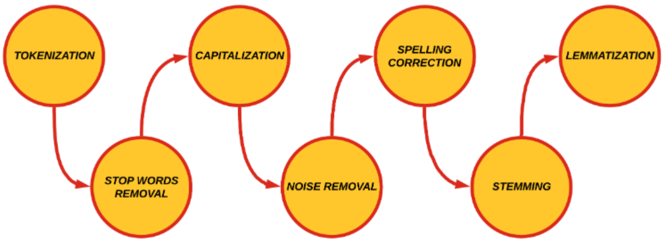


Figure 3: Techniques of data preprocessing phase.

Emphasizing the importance of pre-processing data, in Figure 3, are depicted a number of processes that are followed to clear the data and prepare it for further processing [9]. Such processes as:

- Tokenization - is the process of separating a piece of text into smaller units called tokens. The way the token isformed is based on a delimiter, which in most cases is space. Also, tokens can be words or sub-words, but alsoat a lower level, based on characters.
- Stop Words - are words that are commonly used in one language, that are unnecessary in the data processingpart, and in most cases are ignored because they take up more space in the database, and affect longerprocessing times. In English stop words are words like: "a", "the", "an", "it", "in", "because", "what", to namea few.
- Capitalization - is the part where it is necessary to identify the correct capitalization of the word, where thefirst word in the sentence will be automatically capitalized first.
- Noise Removal - is the process of removing characters, numbers, and parts of text that affect your analysis.These characters can be some special characters, punctuation, source code removal, HTML code removal,unique characters that represent a particular word, numbers, and many other identifiers.
- Spelling Correction - is a problem where the meaning of a particular word can be mispronounced, where theword loses its meaning. This problem can be solved in two ways: with edit distance and another with overlapusing k-gram.
- Stemming - is a process where more morphological variants are produced than the base word or the so-calledroot word. For example different morphological variants of root words "like" such as "likes", "liked", "liking"and "likely".
- Lemmatization - in this technique words are replaced with root words or words that have a similar meaning,and such words are called lemmas.•Syntactic Word Representation (such as N-Gram) - is a contiguous sequence of n items from one part of thetext.• Syntactic N-Gram - are n-grams that are constructed using paths in syntactic trees.–Weighted Words (such as TF and TF*IDF)–Word Embedding (such as Word2Vec, GloVe, FastText)

After finalizing with data pre-processing step, we continue with Dimension Reductions. With dimension reductions we transform the data from a high-dimensional space to a low-dimensional space. The reason for this is that we strive to improve performance, speedup time, and reduce memory complexity. As shown in Figure 4, there are a number of algorithms or techniques in this step that could be implemented, such as: (i) Principal Component Analysis (PCA), (ii) Non-negative Matrix Factorization (NMF), (iii) Linear Discriminant Analysis (LDA)and (iv) Kernel PCA.

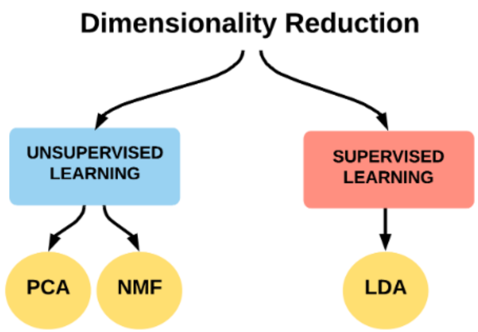


Figure 4: Categorization of dimension reductions algorithms

As part of four-phase model of a text classification system depicted in Figure 2, in the pre-final phase we deal with classifier selection. One of the main concerns is to choose the right classifier model that will be able to perform with a certain set of data to achieve the desired results. Choosing the right classifier model is not an easy task, and is a challenge that is also referred to in the literature as the Algorithm Selection Problem (ASP). Every day we come across applications that use classification algorithms in some hands. The results of the task depend on choosing the right algorithm that will complete a particular job while showing very good performance and problem optimization. In general, there is no single algorithm that can work for every type of problem, and that can learn all the tasks while still being efficient, and this phenomenon is also known as performance complementary [10]. Many factors affect the performance of a particular algorithm, some of which is the amount of data assigned to it for testing and training, the operating system to be executed, the specifications of the machine on which the algorithm will be performed, and many other factors that directly or indirectly affect the selection of the algorithm. Some of the algorithms used for text classification are: Logistic Regression, Naive Bayes, K-Nearest Neighbor (KNN), Support Vector Machines (SVM), Decision Trees, Random Forests, Neural Network algorithms (such as DNN, CNN, RNN) and Combination Techniques. In our experiment we have used K-Nearest Neighbor (KNN) from the conventional models whereas LSTM recurrent neural network from the deep learning models. To conclude, the evaluation phase is encountered as the final step when creating a model for text classification is the evaluation phase. In this phase, algorithms are analyzed or scored to assess how efficiently they performed. Today, the various technologies available today have drastically improved the way people try to gain new knowledge. Technology has greatly influenced the improvement of this process, and at the same time contributed to the development of systems that enable a more efficient and easier learning process. With this fact the use of various Massive Open Online Courses (MOOCs) begins to increase, which bring with them various opportunities, but also challenges. Attempts to identify and analyze the opportunities and challenges of MOOCs both from pedagogical and business standpoint have led to understand how some of the very well known and successful platforms like Coursera, edX and Udacity have contributed to the improvement of their business model through various aspects, using the models for: certification model, freemium model, advertising model, job-matching model, and subcontractor model [8]. During the analysis of these platforms, the authors in [9] concluded that quite a low number of students actually take assessment exams at the end of a MOOC which makes it difficult to assess whether students joining a MOOC are actually learning the content, and hence whether the MOOC is achieving its goal. One of the main components of these platforms is Learning Objects (LOs). Various techniques regarding Learning Objects (LOs) representation are presented, in which it contains pedagogical values [10].

Using the representation features of Learning Objects will provide possibilities to personalize and customizable contents when presenting to learners along with the ability to choose an individual learning path that best suits them, aiming to maximize the learning outcome as claimed in [10]. There are plenty of examples where K-Means, Decision Trees, Deep Neural Network (DNN) and other machine learning techniques have been used for classification purposes [11]. As eLearning platforms are becoming more accessible, where their main goal is to provide a smarter way of learning. The new paradigm of e-Learning also known as Cloud eLearning aiming to offer personalised learning using Cloud resources, where the main challenge is the process of content classification and matching it with learners preferences. As part of this work, the author [12] integrated as middle layer the recommendation systems using hierarchical clustering technique to recommend learners courses or materials that are similar to their needs before proposing a learning path using artificial intelligent automated planner. Also, paper [13] contributes to the classification systems in pedagogical content, with the main focus on the content classification of video lectures. The authors recommended model for the visual content classification system (VCCS) for multimedia lecture videos is to classify the content displayed on the blackboard. Through this recommended model, the authors showed over several stages how lecture videos are processed and then with a combination of support vector machines (SVM) and optical character recognition (OCR) classifies visual content, text and equations [13]. Furthermore in [14], researchers presented the classification and organization of pedagogical documents using domain ontology.

In one of the previous studies [15], the authors of this paper presented a technique for automatic classification of MOOC videos, where the first step is to extract transcripts from video

and then convert them into image representation using a statistical co-occurrence transform. After that, a CNN model with a dataset was implemented which was collected from Khan Academy with a total of 2545 videos, in order to evaluate the technique presented in the paper. Based on label accuracy, the best results were achieved with the CNN model, with the value of 97.87%. Also, similar work has been carried out in [16] where they have proposed a video classification framework, consisting of three main modules: pre-processing, transcript representation, and classifier. In this paper, it was concluded that much better classification results were achieved with general-level than with specific-level, argued with the fact of class overlap that the specific-level category contains.
This paper aims to classify the pedagogical content using two different algorithms, K-Nearest Neighbour as a conventional machine learning model and Long short-term memory (LSTM) as an artificial recurrent neural network architecture used in deep learning.

III. METHODOLOGY

In this section is given the methodology used during the research and the experimental part. Initially a brief introduction regarding the dataset is given, and continuing with explanation of the architectures that are modelled to classify pedagogical content.
Python technology is used for the whole experiment, and specifically to implement the KNN model is used the built-in functions and modules of scikit-learn library, whereas for the implementation of the RNN model is used Keras library, that runs on top of Tensorflow. In the following subsections, the used dataset as part of this experiment is described in detail, following with both models, the KNN and LSTM.

A. *Dataset*

The process of collecting and reviewing data is not an easy task, and in most cases requires a lot of research and finding relevant data that are used to achieve the desired results. The dataset [17] used in this paper for the experimental purposes is used in [16] and it is modelled from scratch. This dataset consists of a total of 12,032 videos collected from the Coursera platform from more than 200 different courses. Coursera categorizes courses into a 2-level hierarchical structure from general level to fine-grained level. The general level consists of 8 categories, the specific level of 40 categories, and the course level of a total of 200 categories. In addition to these three levels that made up the course, a video lesson transcript was also included.
Figure 5 presents the top five most frequent categories, while Figure 6 presents the top five least frequent categories by the number of transcripts that these categories contained. In order for the data to be in the correct format for further analysis and modeling process, the data needs to go under pre-process phase, by preparing, cleaning, and transformed in a desired shape.
The data preparation and preprocessing part depends on the given dataset, and in our case the first step after the review is to remove the noisy data (such as '[MUSIC]' which are recorded very frequently in all transcript records). Following the steps depicted in Figure 3, the entire textual content of the transcript is converted into lowercase, and removed the nonletters characters. Further, the stopwords are removed from the transcript where it helped us reduce the derived words to their particular word stem or root word.

The dataset is transformed finally into the desired shape after finishing the lemmatization process, and it is ready to be used for both architectures that we have modelled, KNN and LSTM described further in the following subsections.

B. *K-Nearest Neighbour model*

K-Nearest Neighbors (KNN) is one of the techniques that is used in both classification and regression. It is known that KNN has no model other than collecting the entire dataset, and there is no need for learning. The predictions made with the KNN for the new data point are by searching the entire dataset for the K most similar instance (so-called neighbors) in relation to the output variant of the K instance [18].

There are a number of steps that the KNN algorithm goes through, such as:

1) *Modify K with the number of specific neighbors.*
2) *Calculate the distance between the available raw data examples.*
3) Sort the calculated distances.
4) Get the labels of top K entries.
5) Generated prediction results for the test case.

In this experiment, while implementing the KNN model, immediately after the process of cleaning and preparing data, is built a dictionary of features, which transforms documents to feature vectors and convert the transcripts of documents to a matrix of token counts using CountVectorizer method.

Then, the count matrix is transformed to a normalized tf-idf representation using Tfidf transformer method. After this is identified the exact number of neighbors which in our case resulted in 7 neighbors. To train the classifier, the dataset is divided into two subsets: 80% for training and 20% for testing. Where the latter subset is used to predict the category for each input text record.

C. Long short-term memory model

Recurrent Neural Networks (RNN) are types of artificial neural networks that allow previous outputs to be used as inputs while having hidden states [19]. These algorithms are mostly used in fields such as: Natural Language Processing (NLP), Speech Recognition, Robot Control, Machine Translation, Music Composition, Grammar Learning, and many others. Typically, a feedforward network maps one input to one output. But as such, the inputs and outputs of neural networks can vary in the length and type of networks used for different examples and applications [20].

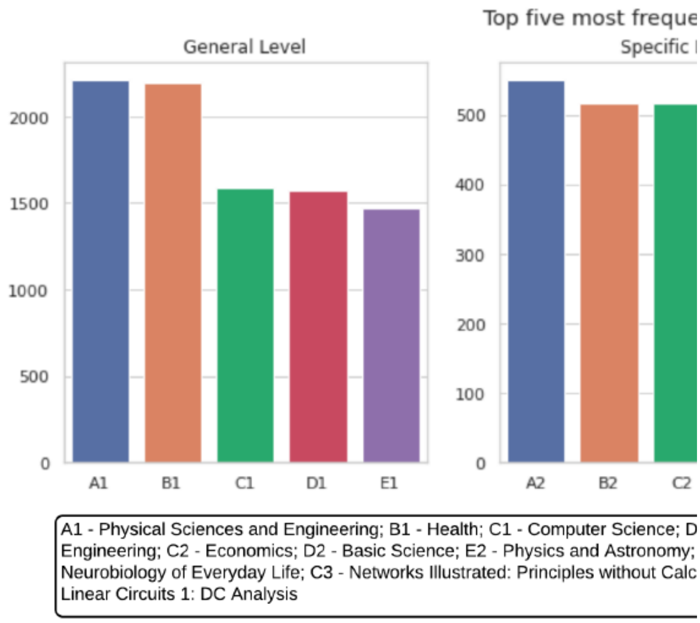


Figure 5: Top five most frequent categories for all three levels.

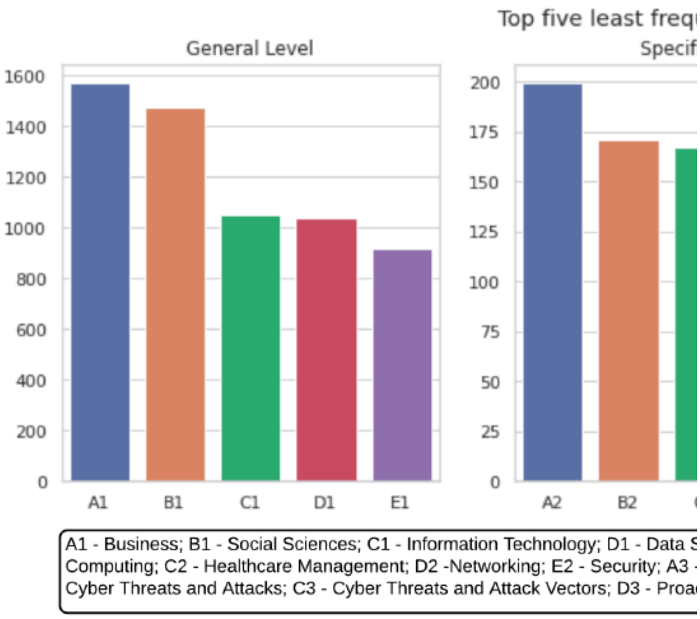


Figure 6: Top five least frequent categories for all three levels

In the implementation of our LSTM model, in order to implement the RNN model, we used the LSTM architecture that remembers values over arbitrary intervals. As part of this architecture firstly are created sequence models as the input layer to our network, then adding the Embedding layer which encodes to integer values the textual data entered as input, and as a result of this layer each word is then represented by a unique integer. For this layer, we have specified three required parameters with their respected values:

- Maximum number of words - which in our case is 50000.
- Embedding Dim - 100.
- Input length - shape of X value which for us is 3002.

Further are dropped out hidden and visible units between the layers in the network, with a dropout rate of 0.2, the same value is for recurrent dropout as well. This is followed by the implementation of LSTM layer, and Dense layer to which we passed as the first parameter the number of units denoting the dimensionality of the output space, which in our case depends on the number of categories that are selected to classify, and as the second parameter the activation function, in this case is chosen the softmax function. And as a final step, is used categorical crossentropy as a loss function, and Adam as an optimizer of the network. To prevent underfitting or overfitting of the network, and to select the appropriate number of training epochs is used EarlyStopping with 'val loss' as a monitoring metric with patience of 3 epochs.

IV. RESULTS

Table I shows the classification results with the conventional model using K-Nearest Neighbours algorithm. As shown in Table I, the general level based on the precision metric has shown a very good result, 92.63% of accuracy whereas 87.89% accuracy is estimated by precision metrics specific level. And at the course level, also based on the precision metric reaches 78.59%. Analyzing the results for all three levels, we notice that the percentage of accuracy decreases from the upper level (general level) up to the lower level (course level). In our case, the general level consists of 8 sub-categories, the specific level of 40 sub-categories, and the course level consists of 200 subcategories.

From this we can infer that that the number of subcategories for a single level by which the video is classified on the Coursera platform differs in each level.

TABLE I. CLASSIFICATION RESULTS WITH K-NEAREST NEIGHBOURS

Category	***Precision (%)***	***Recall (%)***	***F1 Score (%)***	***Accuracy (%)***
General Level	92.63	92.52	92.53	92.52
Specific Level	87.89	87.58	87.49	87.58
Course Level	78.59	76.73	76.11	76.73

Table II shows the classification results with the Recurrent Neural Networks, more specifically with an Long Short-Term Memory (LSTM) architecture. Using LSTM classifier, the general level based on the precision metric reaches 88.22% of accuracy whereas in the specific level, 72.31%. Finally, at the course level, the results shows 59.49% of accuracy. Analyzing the results using LSTM architecture the highest accuracy is achieved at the general level, followed by a specific level, while the lowest accuracy is achieved at the course level.

TABLE II. CLASSIFICATION RESULTS WITH RECURRENT NEURAL NETWORKS

Category	***Precision (%)***	***Recall (%)***	***F1 Score (%)***	***Accuracy (%)***
General Level	88.22	87.71	87.68	87.71
Specific Level	72.31	69.93	70.13	69.93
Course Level	59.49	52.91	53.99	52.91

V. CONCLUSION AND FUTURE WORK

In this paper are presented and discussed the classification results of the conducted experiment for all three category levels (General, Specific and Course level) using both architectures, KNN and LSTM. We can conclude that better results are achieved for levels with a smaller number of categories than for levels with a larger number of categories. In our case, as the category number increased in classes the results decreased. With this, we claim that the classification results are directly affected by the number of categories that each level contains. From results shown in Table I and Table II KNN reached 92.52% of accuracy compared to LSTM with 87.71% at general level, 87.58% compared to 69.93% at specific level and finally 76.73% compared to 52.91% at course level. The conducted results could be affected from several factors. First, the quantity of data required for LSTM, since a large number of categories increases the complexity of the problem, and thus requires more data to train the model. The result could have been affected due to the high similarity of different transcripts. Many of the transcripts belonged to different classes at the course level, and they had many similarities in the context of the sentences and keywords, so the model could not properly distinguish in which class the transcripts belonged. However, the final results gives us a spark for future work to investigate more on recurrent neural networks like, applying hyperparameters tuning, or even expand the number of architectures to further investigate the pedagogical content classification.

REFERENCES

[1] Fabrizio Sebastiani. Machine learning in automatedtext categorization.ACM computing surveys(CSUR), 34(1):1–47, 2002.

[2] Cicero Dos Santos and Maira Gatti. Deep convo-lutional neural networks for sentiment analysis ofshort texts. InProceedings of COLING 2014,

the25th International Conference on ComputationalLinguistics: Technical Papers, pages 69–78, 2014.

[3] Zenun Kastrati, Ali Shariq Imran,and Sule Yildirim Yayilgan. A general frameworkfor text document classification using semconand acvsr. InInternational Conference on Human Interface and the Management of Information,pages 310–319. Springer, 2015.

[4] Arfinda Ilmania, Samuel Cahyawijaya, Ayu Pur-warianti, et al. Aspect detection and sentimentclassification using deep neural network for in-donesian aspect-based sentiment analysis.In2018 International Conference on Asian LanguageProcessing (IALP), pages 62–67. IEEE, 2018.

[5] Ali Shariq Imran, Sher Muhammad Daudpota,Zenun Kastrati, and Rakhi Batra. Cross-culturalpolarity and emotion detection using sentimentanalysis and deep learning on COVID-19 relatedtweets. IEEE Access, 8:181074–181090, 2020.

[6] Zenun Kastrati, Ali Shariq Imran, and Arianit Kurti.Weakly supervised framework for aspect-basedsentiment analysis on students' reviews of MOOCs.IEEE Access, 8:106799–106810, 2020.

[7] Alya Itani, Laurent Brisson, and Serge Garlatti.Understanding learner's drop-out in MOOCs. Ininternational conference on intelligent data engi-neering and automated learning, pages 233–244.Springer, 2018.

[8] Hannes Max Hapke, Hobson Lane, and ColeHoward. Natural language processing in action,2019.

[9] Kamran Kowsari, Kiana Jafari Meimandi, MojtabaHeidarysafa, Sanjana Mendu, Laura Barnes, andDonald Brown. Text classification algorithms: Asurvey. Information, 10(4):150, 2019. [10] Irfan Khan, Xianchao Zhang, Mobashar Rehman,and Rahman Ali. A literature survey and empiricalstudy of meta-learning for classifier selection.IEEEAccess, 8:10262–10281, 2020.

[10] Jake Lever, Martin Krzywinski, and Naomi Altman.Erratum: Corrigendum: Classification evaluation.Nature Methods, 13(10):890–890, 2016.

[11] Fisnik Dalipi, Sule Yayilgan, Ali Shariq Imran, andZenun Kastrati. Towards understanding the MOOCtrend: pedagogical challenges and business oppor-tunities. InInternational Conference on Learning and d Collaboration Technologies, pages 281–291.Springer, 2016.

[12] Krenare Pireva, Ali Shariq Imran, and FisnikDalipi.User behaviour analysis on LMS andMOOC. In2015 IEEE Conference on e-Learning,e-Management and e-Services (IC3e), pages 21–26.IEEE, 2015.

[13] Fisnik Dalipi, Ali Shariq Imran, and Zenun Kastrati.MOOC dropout prediction using machine learningtechniques: Review and research challenges. In2018 IEEE Global Engineering Education Confer-ence (EDUCON), pages 1007–1014. IEEE, 2018.

[14] Krenare Pireva and Petros Kefalas. A recommendersystem based on hierarchical clustering for cloude-learning. InInternational Symposium on Intel-ligent and Distributed Computing, pages 235–245.Springer, 2017.

[15] Ali Shariq Imran and Faouzi Alaya Cheikh. Black-board content classification for lecture videos. In2011 18th IEEE International Conference on ImageProcessing, pages 2989–2992. IEEE, 2011.

[16] Ali Shariq Imran and Zenun Kastrati. Pedagogicaldocument classification and organization usingdomain ontology. InInternational Conference onLearning and Collaboration Technologies, pages499–509. Springer, 2016.

[17] Houssem Chatbri, Kevin McGuinness, SuzanneLittle, Jiang Zhou, Keisuke Kameyama, Paul Kwan,and Noel E O'Connor. Automatic mooc videoclassification using transcript features and convolu-tional neural networks. InProceedings of the 2017ACM Workshop on Multimedia-based Educationaland Knowledge Technologies for Personalized andSocial Online Training, pages 21–26, 2017.

[18] Zenun Kastrati, Ali Shariq Imran, and ArianitKurti. Integrating word embeddings and documenttopics with deep learning in a video classificationframework.Pattern Recognition Letters, 128:85–92, 2019.[

[19] Zenun Kastrati, Arianit Kurti, and Ali Shariq Imran.Wet: Word embedding-topic distribution vectorsfor mooc video lectures dataset.Data in brief,28:105090, 2020.

[20] Jason Brownlee.Master Machine Learning Al-gorithms: discover how they work and implementthem from scratch. Machine Learning Mastery,2016.

[21] Afshine Amidi and Shervine Amidi. Vip cheatsheet:Recurrent neural networks, 2018.

[22] Larry Medsker and Lakhmi C Jain.Recurrentneural networks: design and applications. CRCpress, 1999

Potential of Use of the Republic of Serbia Renewable Energy Sources

gvozdena2001@yahoo.com

Nenad A. Marković
Academy of Applied Studies of Kosovo and Metohija,
Department Uroševac – Leposavić, Serbia
nen.mark74@yahoo.com

Nenad T. Stanojević
Academy of Applied Studies of Kosovo and Metohija,
Department Uroševac – Leposavić, Serbia
nenads25@gmail.com

Jelena D. Rajović
Academy of Applied Studies of Kosovo and Metohija,
Department Uroševac – Leposavić, Serbia

Abstract—**Renewable energy sources are energy sources that can be completely or partially renewable. Their use is lately more and more prominent because of the increasing pollution of the planet Earth. Support for the sustainable development of renewable energy sources has become one of the main goals of the European Union, and thus of the Republic of Serbia. The Republic of Serbia has a significant energy potential that it will strive to use in the future. Therefore, in this paper we will present the potential of renewable energy sources of the Republic of Serbia, as well as the possible potential of their use in electricity production.**

Keywords-Renewable energy sources; The Republic of Serbia; Energy potentials; European Union Directive

I. Introduction

Renewable energy sources are one of the most efficient solutions for clean energy and sustainable development. Throughout the world today, the strategic positioning of states, nations and companies for access to the remaining natural resources, especially mineral energy sources such as oil and gas, but also the technological competition in energy efficiency and commercial use of renewable energy sources, is still in progress [1].

Renewable energy sources are the priority of energy union. In the countries of the European Union, and thus the Republic of Serbia, renewable energy sources are one of the most current issues as possible producers of electricity.

Directive 2001/77/EC [2] on promotion of electricity produced from renewable energy sources and Directive 2003/30/EC [3] on promotion of the use of biofuels and other renewable fuels in transport are among the first directives on renewable resources.

Directive 2009/28/EC on renewable energy has set a mandatory target of 20% of the total share of renewable energy sources in electricity consumption by 2020 [4].

In 2018, the revised Directive on Renewable Sources (Directive (EU) 2018/2001) entered into force, as part of the "Clean Energy for All Europeans" package. The aim of this directive is that the total share of renewable energy sources in electricity consumption will be 32% by 2030 [5].

The Republic of Serbia in 2007 ratified the Kyoto Protocol to the United Nations Framework Convention on Climate Change. The Kyoto Protocol aimed to reduce the emission of harmful gases that cause the "greenhouse" effect, thus reducing the rate of warming of the atmosphere. By ratifying this protocol, the Republic of Serbia has given a clear signal to its commitment to renewable energy sources [6], and gained the status of a net seller of emission credits and the right to finance projects to improve energy efficiency in the country. This lead to the increase of competitiveness of its economy on the world market.

II. National Strategy of the Republic of Serbia on Electricity Production from Renewable Energy Sources

Ministry of Energy, Development and Environmental Protection of the Republic of Serbia in 2013 adopted the National Action Plan for the use of renewable energy sources. [7] This Action Plan is the result of the international obligation of the Republic of Serbia from 2006: "Law on the Ratification of the Treaty establishing the Energy Community between European Union and The Republic of Albania, the Republic of Bulgaria, Bosnia and Hercegovina, the Republic of Croatia, The Former Yugoslav Republic of Macedonia, the Republic of Montenegro, Romania, the Republic of Serbia and United Nations Interim Mission in Kosovo in accordance with the resolution 1244 of the United Nations Security Council" ("Official Gazette of the Republic of Serbia" No. 62/06 [8]). Action Plan defines fundamental goals of the energy policy of the Republic of Serbia:

- Development of the energy infrastructure.

- Diversification of energy sources to ensure security of supply.
- Introduction of modern technologies in energy sector (especially technologies that will encourage the economic development of the country).
- Reduction of the growth of final energy consumption.
- Increasing the energy efficiency.
- Increasing the use of renewable energy resources [7].

In National Strategy of Sustainable Development (2008) sector for renewable energy sources, the following goals have been defined [9]:

- Extensive research on potential of sustainable energy sources.
- Determining the technology for which the introduction of incentive measures and mechanisms is justified.
- Adoption of regulations (tax deductions, incentive prices, etc.) for stimulating the use of renewable energy sources.
- Increasing the scope of use of renewable energy sources.
- Education and awareness raising as an incentive for the inclusion in the production and the use of energy form renewable sources.

The Republic of Serbia has significant renewable energy resources to meet these strategies, and the Government of the Republic of Serbia has developed an energy development strategy for the period to 2025, with projections up to 2030, Fig. 1.

Figure 1. Development of the Republic of Serbia strategy for the development of energetics to 2025, with projections up to 2030 [10].

III. POTENTIAL OF RENEWABLE ENERGY SOURCES IN THE REPUBLIC OF SERBIA

Access to energy from renewable sources is one of the most important long-term policy decisions a country can make. Its development brings a number of benefits. One of the main benefits is that renewable energy technologies in order to achieve sustainable development provide the opportunity to reduce carbon dioxide in the atmosphere, and thus slow down climate change [11].

Energy efficiency and introduction of renewable energy sources in production, transmission, distribution and consumption/satisfaction of energy needs in energetic sector are the most important mechanisms for combating climate changes [12].

The studies have shown that energy potential of the renewable energy sources can satisfy about 25% of electric power market needs in the Republic of Serbia [12]. In the future, it is predicted that the mainstay of energy independence of the Republic of Serbia will be renewable energy sources. However, the economic viability of renewable energy sources is only at the level of estimates. It is estimated that information on the cost-effectiveness of using renewable energy sources in the Republic of Serbia is very limited and needs to be confirmed.

Energy produced from non-fossil renewable sources represents renewable energy sources, including hydropower, solar energy, wind energy, geothermal energy, ocean energy, bioenergy. Some of the most significant are wind energy, solar energy, bioenergy and hydropower.

Regarding the wind energy, it has been estimated that the technically usable potential in the Republic of Serbia is 2300 GWh and that it can replace about 2% of total consumption of the electric energy. Existing wind data suggest that at several locations there are wind speeds between 6-7 m/s, while foreign experience suggests that this corresponds to an annual load of about 18-25%. Only the location of the top of Stara Planina, Midzor, with an established annual average of 7.6 m/s, can be classified as good, with an expected load of about 28% [13].

The Republic of Serbia has significant possibilities in the area of biomass combustions. It has been estimated that biomass energy potential of the Republic of Serbia is about 3.405 million ten (1 ten=1 tonof equivalent oil =11630 kWh =11.63 MWh). It consists of residues in the wood industry (1.53 million ten), the potential of agricultural biomass (1.67 million ten residues in farming, livestock, fruit growing, viticulture and primary fruit processing), while the potential of biodegradable municipal waste is estimated at 205 thousand tons. Biodegradable waste (except municipal waste) also includes waste edible oils and waste of animal origin (rendering slaughterhouse waste) in the total amount of 0.043 million ten/year [1].

A large part of the Serbian economy is based on agricultural production and agriculturally oriented industry. The northern part of the Republic of Serbia, the Province of Vojvodina and the territories along the Sava and Danube rivers are the main areas with the sources of biomass waste [14].

Solar energy is one of the energy potentials of the Republic of Serbia. Number of hours of solar radiation over most of the territory of the Republic of Serbia is much higher than in many European countries and is over 2000 hours. The average solar radiation is about 40% higher than the European average, but the use of this energy in production is still far behind the countries of the European Union. Areas in the Republic of Serbia in which a large number of hours of sunshine are registered and the annual ratio of actual radiation and total possibilities make up approximately 50% of the territory.

Fig. 2 represents the annual average of day energy of global radiation on the horizontal surface.

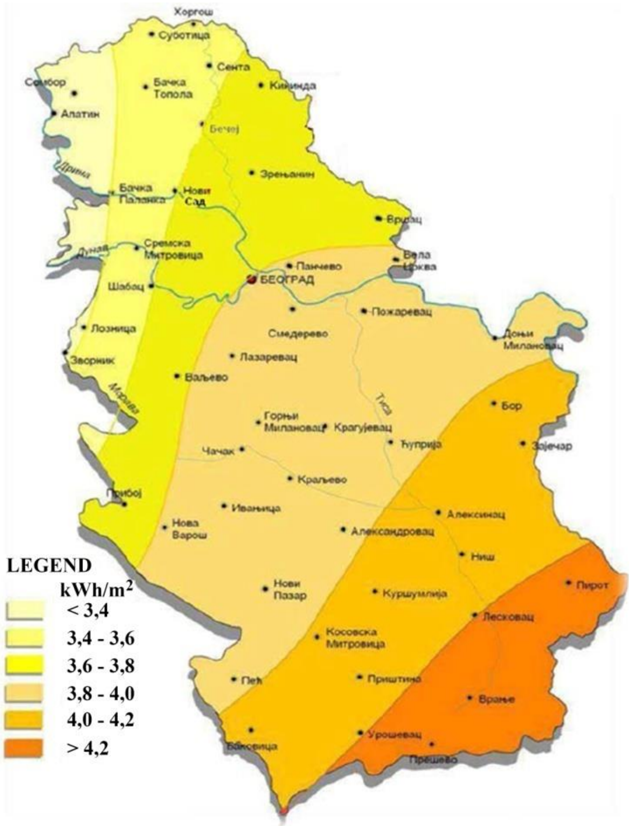


Figure 2. Annual average of day energy of global radiation on the horizontal surface [14].

There are different ways of using solar energy: heating with solar energy and converting solar energy into electricity, which is done in two ways, by concentrating solar energy and photovoltaic collectors (panels). An advanced method is the direct production of electricity by photovoltaic panels [15].

In addition, the Republic of Serbia has a rich geothermal potential of over 2300 GWh per year. The use of geothermal energy encourages the importance of developing spa, health and recreational tourism. The most developed geothermal areas are in Vojvodina, and spa towns Vranjska, Sijerinjska and Jošanička Banja have the highest water temperature.

The draft strategy for the development of energy in the Republic of Serbia for the period to 2025, with projections until 2030, states that the potentials of renewable energy sources in the Republic of Serbia are significant and are estimated at 5.65 million ten per year.Of this amount, more than 60% is biomass potential, the use of which is currently estimated at about 30% of the available potentials. The available technical hydro potential participates with about 30% in the total potentials of renewable energy sources. Half of this amount has been already been used. Regarding the other renewable sources, currently only the use of geothermal energy is partially monitored and balanced [1].

On Fig. 3 the projection of capacity building for electric energy production using renewable energy sources is presented.

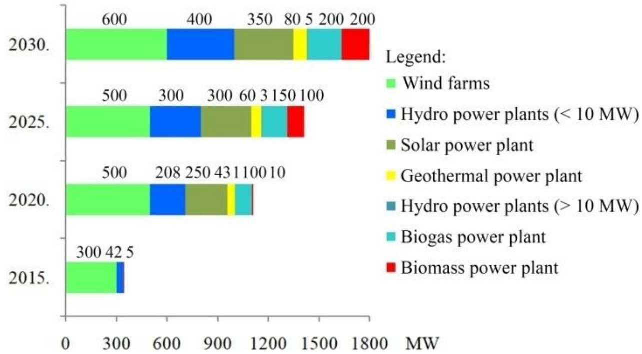


Figure 3. Projection of capacity building for electric energy production using renewable energy sources [1].

Legal framework for the use of renewable energy sources in the Republic of Serbia is reflected in the fact that Serbia, in the process of joining the European Union will shape its legal regulations on renewable energy resources in accordance with European Union legal regulations.

It is interesting to note that the funding of renewable energy sources is one of the crucial factors in the development of any country. The challenge is to provide appropriate financial instruments that will enable renewable energy sources to become common and to maximize their potential in the market [16].

IV. CONCLUSION

In recent years, renewable energy sources have played an increasing role in energy production. Various environmental constraints, such as limiting carbon dioxide emissions from electricity generation to prevent the greenhouse effect, strongly emphasize the benefits of renewable energy sources. One of the strategic goals of the State should be to support the construction of new capacities for the utilization of renewable energy sources. Renewable energy production should also be increased to reduce import dependence and increase energy security.

At the end of this presentation, we can conclude that the Republic of Serbia has significant energy potential, but unfortunately not sufficiently exploited.

REFERENCES

[1] Republic of Serbia, Ministry of Energy, Development and Environmental Protection, Development of the Energy Development Strategy of the Republic of Serbia for the period to 2025 with projections till 2030, Phase II, University of Belgrade, Faculty of Mining and Geology, Belgrade, June 14, 2013.

[2] https://eur-lex.europa.eu/eli/dir/2001/77/oj, downloaded, 05.04.2021.

[3] https://eur-lex.europa.eu/eli/dir/2003/30/oj, downloaded, 05.04.2021.

[4] https://eur-lex.europa.eu/eli/dir/2009/28/oj, downloaded, 05.04.2021.

[5] https://eur-lex.europa.eu/eli/dir/2018/2001/oj, downloaded, 05.04.2021.

[6] Government of the Republic of Serbia in the Service of Citizens, Belgrade, September 14, 2007, https://www.srbija.gov.rs/vest/73811/ratifikacija-kjoto-protokola-od-vitalnog-znacaja-za-srbiju.php, downloaded 06.04.2021.

[7] The Republic of Serbia, Ministry of Energy, Development and Environmental Protection, National Action Plan for the Use of Renewable Energy Sources of the Republic of Serbia, Belgrade, 2013.

[8] Republic of Serbia, “Official Gazette of the Republic of Serbia,” No. 62/06.

[9] Republic of Serbia, The National Strategy of sustainable development 2008-2017, “Official Gazette of the Republic of Serbia,” No. 57/08.

[10] M. Banjac, “Renewable energy sources in Serbia, legal framework, obligations and planes,” SREM regional development agenda, 6. December 2012.

[11] Renewable Energy, Guide for Parliamentarians, United Nations Development Programme, 2013.

[12] Renewable Energy Sources in Serbia, Recommendations, Potentials and Criteria, Center for Ecology and Sustainable Development, Project “Towards sustainable energy in South East Europe,” CEE Bankwatch Network in Serbia, Subotica, December 2008.

[13] D. Kragulj, M. Parežanin, “Renewable energy sources as a significant component of sustainable development,” VIII Conference of Businessmen and Scientists, SPIN, Faculty of Organizational Sciences, Belgrade, 01-02. November 2011.

[14] Ž. Despotović, “Renewable energy sources - state and perspectives in the world and in Serbia,” Belgrade, April 2012.

[15] M. Radaković, “Renewable Energy Sources 1,” Solar Energy Society “SRBIJA SOLAR”, Zrenjanin, 2006.

[16] N. Sandić, “Renewable energy sources as a factor of the economic development of Serbia,” FEFA-Faculty of Economics, Finance and Administration, Singidunum University, Belgrade, 2013.