International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (Mon-20..)

Oasis: Mobile Based Comprehensive Student Support System for Undergraduates

S.M.S Fernando[1], P.M.D.C Rajapaksha[1], M.P Gamage[2], S.H.U Madushan[1], K.D.L.J Karunathilaka[1]

[1] *Department of Software Engineering Sri Lanka Institute of Information Technology, Colombo, Sri Lanka*
[2] *Department of Information Technology Sri Lanka Institute of Information Technology, Colombo, Sri Lanka*

E-mail address: it20649012@my.sliit.lk, it20644062@my.sliit.lk, manori.g@sliit.lk, it20646288@my.sliit.lk, it20646806@my.sliit.lk

Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: Current higher educational institutions face significant challenges in meeting the demand for on-site student support services due to increasing student enrollment and budget constraints. Advancements in technology offer promising solutions through support systems that could be made available online. This research aims to design a readily available Comprehensive Student Support System for undergraduate students, which provides personalized support in multiple facets. Main objectives of the study are to provide proactive mental health support, enable career awareness and readiness, improve academic participation out of the course-curriculum, and to identify the best tools and technologies for system implementation. The system's main components include proactive support for mental well-being through the practice of diary writing, job recommendations based on the resume data to help students identify their current job placement status, career path guidance that supports students to discover new avenues available based on their interests, and facilitating networking opportunities for professional growth through academic related external event recommendation. Utilizing advanced methods in Natural Language Processing and Machine Learning has proven effective in achieving the objectives. The system is evaluated using Functional and User Acceptance Testing methods. Key insights from the system's evaluation highlight its usability and acceptance as a student support system.

Keywords: Student Support Services, Student Support Systems, Mental Well-being, Natural Language Processing, Machine Learning, Sentiment Analysis, Resume Parsing, Career Guidance, Personalized Study Support.

1. INTRODUCTION

A. Background

The stroll through undergraduate life is a very crucial moment of an individual's life that is marked by opportunities as well as emotional shifts, challenges, and stress due to unforeseen fear of prospects. Nevertheless, getting through these intricate situations is challenging for most students, which may result in various problems, including emotional stress, difficulties in academic involvement, confusion about the future career prospects, and uncertainty and fear of stepping into professional world. Due to the growing need for mental wellbeing and the necessity for career guidance and academic support, Higher Educational Institutes (HEI) are trying their best to implement approaches to support the students [1], [2], [3], [4]. Traditional on-site Student Support Services available in HEI for the purpose, face significant challenges that can impede their effectiveness and prevent students from fully benefiting from available resources. Key issues include reluctance of students to seek face-to-face support, limited accessibility for remote or part-time students, long wait times for appointments that may conflict with academic or work schedules, and the digital dependance of modern undergraduates.

In such regard, digital solutions, especially mobile applications, represent an attractive solution for the provision of tailored, reachable, and effective support for students. The implementation of online Student Support Systems (SSS) in HEI presents a viable solution to overcome the limitations of traditional on-site support services. These online platforms offer unparalleled accessibility and convenience, enabling students to access a range of support services from any location and at any time, using mobile devices or computers. This flexibility is especially advantageous for the new generation who depend on digital devices, working professionals, or those with disabilities, who often encounter barriers when

E-mail:author's email

attempting to utilize on-campus services. By harnessing the power of technology, these systems can deliver personalized recommendations and proactive interventions, fostering increased student engagement and academic success. The adoption of online SSS presents a promising opportunity for HEIs to address the shortcomings of traditional SSS and enhance the overall student experience [5], [6], [7], [8], [9].

Considering these authors identified an urgent call for a comprehensive SSS for undergraduates that supports their well-being, academic success, and confidence in facing the professional world. An overview of the proposed system is given in the following section.

B. Components of the Proposed System

The mobile-based comprehensive SSS named 'Oasis' proposed in this study addresses the holistic requirements of undergraduate students in a personalized manner. It comprises four main components, each targeting a specific aspect of the undergraduate experience:

- Proactive support for Mental Well-being: In this component, students are encouraged to write a diary entry for the day with emphasis on the happenings of the day, that may aid self-reflection and inner realization. The system determines the mood communicated through the diary entries and offers instant responses based on the assessed mood. For entries that indicate negative moods, it proposes personalized mood improvement methods. In cases of neutral or positive mood, it gives positive feedback. This open and direct communication helps in the creation of a supportive and empathy based digital atmosphere for students. Also supports identifying the need for any early interventions.
- Job Recommendation based on resume data: This component facilitates a student to upload the resume and is used to comprehend the major, skills, experiences, and the educational qualifications. Based on this data, it proposes job opportunities that best match, helping them to identify career opportunities available for them now based on their qualifications.
- Career Path Guidance: This component provides customizable job opportunities available to match students' major, interests, and core skills. Also, it suggests the additional skills and qualifications the student needs to build to target that specific job. This knowledge facilitates students to identify the possibilities and take the right decision in fulfilling the requirements for their preferred future profession. The goal is to provide career options which are not only compatible with the student's current qualifications, but identifying other available options that best suits the students' academic interests, also recommending skills and qualifications that the student needs to improve in order to conquer that job. This can act as a connection which couples the educational involvements of students of today with their future work profession.
- Networking through Related Event Recommendation: This component provides recommendations of academic events like workshops, guest lectures, conferences, and exhibitions that would be of interest and value to the student. By recommending these extra opportunities which are relevant, this module aims to remove the gap between knowledge and reality, hence, enabling better quality of education and active engagement of students. These event suggestions, which are not just meeting academic curriculum requirements, contribute significantly to identifying networking opportunities in the domain and to face the professional world with confidence.

C. Objectives and Significance of the Research

The aim of the study is to design a multi-faceted SSS for undergraduate students that is easily accessible anytime and anywhere, which incorporates essential areas of their personal and academic growth, integrating the best possible tools and technologies. The objectives of the study are to:

- Provide proactive mental health support through personalized recommendations.
- Enable career awareness and readiness through personalized job and skill recommendations.
- Improve academic participation out of the course-curriculum, with academic related event recommendations, enabling networking opportunities for the growth of students.
- Identify the best tools and technologies to develop the system to provide a better service.

Although the already existing systems are providing support in certain fields, Oasis marks a significant stride towards a holistic and technology driven solution with a high degree of personalization. In addition, this project adds to the growing knowledge base in the areas of educational technology and SSS, and insights into selection of ML and NLP technologies to solve real-world issues; thus, facilitating as a model for further revolution. When

compared to existing undergraduate support systems, Oasis stands out in several key areas as follows.

- Holistic Approach: One concern of previous student support systems is that they address only one dimension of academic life. But this provides a wide range of versatile support to undergraduate students covering various aspects of their needs.

- Self-reflecting e-diary entries: The system includes an e-diary, that offers the students an interactive channel to reflect and share their daily experiences and feelings with the system anonymously. This component aims to create a practice of mindfulness supporting proactive intervention for mental well-being throughout the challenging undergraduate years.

- Personalized career counseling: Comprehensive Undergraduate Support System uses AI and ML to empower its capability to provide users with unique and personalized recommendations on job opportunities. To improve the level of personalization research tried to consider as many student attributes as possible. Within other available support systems, there is less consideration to career counselling.

- Technological Innovation: Providing the support services through a mobile application and the integration of modern technologies specially the paraphrase-minilm-l6-v2 model is a major attraction of Oasis that is different from the other career guidance systems. Such technological complexity allows for a more sophisticated comparison of data leading to accurate search results.

2. LITERATURE REVIEW

This review explores existing research and applications, highlighting the contributions and limitations of current approaches, and identifying gaps that the Comprehensive Undergraduate Support System – 'Oasis' aims to address.

A. Discussion of Previous Work in Similar Domains

Educational technology landscape has seen major developments in the past decade, with researchers and developers exploring numerous approaches to improve the student learning experience and their general well-being. Technology has been integrated into student support services, especially mobile applications, which target a variety of student needs, such as their emotional state, career counseling, and academic involvement. This section provides an overview of previous work, identifying the contributions that have informed the development of Oasis.

Supporting students mental well-being has been one of the focal areas of educational research [10], [11], [12], [13]. Some researchers have looked into the use of mobile applications in mental health monitoring and intervention. Most of these applications usually utilize different self-assessment questionnaires and mood tracking features to detect possible mental health problems and give users appropriate coping mechanisms or referral. Nonetheless, most of the current applications are limited by their use of self-reported data and lack of a more detailed analysis of the users' self-reported Natural Language (NL) inputs that can provide more profound insights into the users' mood. There are many researches done in different domains that utilize NL input analysis techniques to identify the meaning behind the text [14], [15], [16], [17], [18], [19].

Technology has enabled new applications in multiple fields that could automate personalized recommendations to users [20],[21],[22]. In career guidance related research, several studies have used machine learning (ML) approaches in order to recommend suitable jobs by matching job seekers' profiles with suitable job vacancies [23], [24], [25], [26]. These systems use NLP and ML algorithms to analyze resumes with job descriptions thus identifying possible matches. Although successful in providing job suggestions, they offer them job options of what they are already suitable. Existing systems do not offer recommendations of other possible jobs available in the industry that matches with their major and interests for which they could target.

Improving academic engagement as well as out-of-curriculum academic involvement is important specially for undergraduates. There are several systems that focus on improving academic engagement of students [27], [28], [29], [30]. Learning materials are now available on platforms such as Coursera offering a variety of courses and resources to the students. Also there are systems that can predict student performance and suggest personalized course recommendations. However, as these platforms make a big impact in recommending personalized learning environments, there is a gap in combining study support with the professional world. Specially the undergraduates should be given exposure to real-world professionals in the industries of their respective domains.

B. Gap Analysis in Existing Undergraduate Support Systems

Investigation of existing SSS uncovered that many researchers strive to improve student life in multiple aspects such as study support, mental health services, and

career counseling. However, an in-depth analysis reveals that the existing systems are not able to offer the necessary integrated, holistic approach that would cover the various needs of students. This gap analysis identifies some critical areas where the existing systems are underperforming as listed below which stresses the need for a more harmonized solution.

1) Fragmentation of Services: One of the major gaps is the dispersal of support services. The current systems are often siloed, with separate services being offered for emotional wellness, career counselling, and academic support. This gap can result in redundant and awkwardness of students who have to deal with several platforms just to get what they need. The absence of a central platform may lead to under-provision of support to students who could miss out on critical aspects of their development or may face barriers in accessing available resources.

2) Unavailability of Personalized career councelling: Career services usually address resume construction and job placement, whereas academic services are oriented to selection of a major and academic performance. In career counceling fares conducted by HEI they try to fit the students to their available vacancies but it would not be the best match for the student. The gap between the academic curriculum and the selection of a career option leaves students to bridge it by themselves. Another important gap is that the existing support systems have not given importance to present other inspiring job possibilities in their respective study domain that could be targeted if necessary skills are aquired. Career awareness and prepardness is of much importance to the undergarduates due to the high job demand at present. Also it helps them to identify their current job placement status and also awareness of other job opportunities available based on their major and interests so that they have enough time to prepare and achive their dream job.

3) Inadequate Proactive Support for Mental Well-being: Though the traditional on-site counselling services are very important, it is possible some students may not be able to get or may be reluctant to use them. And even though there are many systems available to provide mental health services, the support is approached in a reactive rather than a proactive manner. There are no continuous engagement tools that track students mood over time and offer early intervention. And most systems available utilize self-assessment questionnaires to identify issues and lacks the use of self-reported data for the purpose.

4) Insufficient use of Technology and Service Accessibility: Technology utilization in undergraduate support systems is inefficient. The increasing prevalence of smartphones and the availability of internet access make mobile applications a powerful instrument for rendering support services to students. But many current systems increasingly fail to utilize this potential. This gap of technology highlights the chance to use mobile platform with relevant technology stack for consistent, available, and integrated support.

The Oasis is designed to overcome these shortcomings by providing an integrated system that is accessible anywhere and anytime and equipped with the latest technology to gain full potential of the system.

3. SYSTEM ARCHITECTURE AND DESIGN

At the beginning several consultations are done with the stakeholders of the proposed system, such as students, academic advisers, and career counselors, with the intension of understanding and getting familiar with key features and requirements for such a system. A user-responsive design of the Oasis as given in Figure 01 is realized through modern advanced technologies to meet the diverse requirements of students, fostering their emotional, academic, and professional growth. The following sections detail the architecture of the system and the design of its components highlighting the selection of the technologies.

A. Mobile Application Architecture

The architecture of the mobile application has been developed for it to be scalable, modular, and user-friendly, which guarantees that students can access the full range of support services effortlessly. The application is built on a client-server model with a mobile app serving as the client interface for user interaction and backend server handling data processing, analysis, and content management.

- Client-Side (Mobile Application): The client-side is implemented in Flutter, the widely known open-source UI software development kit developed by Google. From a single codebase, flutter enables the development of natively compiled mobile applications, providing a fast pace and excellent user experience for both iOS and Android platforms. Dart, which is the language of Flutter, has a clear syntax and a powerful set of libraries, which can be used for developing more complicated features.
- Server-Side (Backend Services): The backend services are based on Firebase, a mobile and web application development platform that offers a set of cloud services such as a real-time database, authentication, and hosting. Firebase gives the system capability to securely handle user information, perform the data processing and render custom content to users as it happens. The scalable cloud database can be used to store

various student data such as diary entries, resumes, and personal profiles.

- Communication: The client and server talk using RESTful APIs, which guarantees secure and efficient data communication. This configuration allows the live generation of the application content, interactive user responses, and the embedding of ML models for data processing.

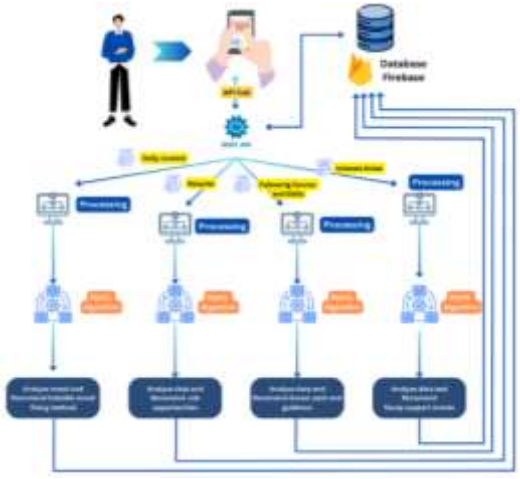

Figure 1. High level system diagram

B. *High-Level Design of the proactive support for mental Well-being*

1) *User Interaction and Data Collection*

Upon launching the Daily Diary tool, users are welcomed by a visible, simple interface that helps them to reflect on their day. The app comprises subtle alerts and inspirational texts that encourage users to stick to diary habits consistently, with the goal of ensuring constant mental health monitoring. The app tries to support an adequate rate of user interaction through gamification methods, such as streaks and badges, rewarding users for every consecutive day of filling in diary entries.

2) *Sentiment Analysis Using the Roberta Model*

In the sphere of mood detection from diary entries, sentiment analysis is one of the key methods in NLP that supports identification of sentiment of written text and contextual relevancy.

RoBERTa (Robustly Optimized BERT Pretraining Approach) model has a better understanding of the text context, which is useful for the diary analysis. RoBERTa enhances BERT's performance through optimizing essential hyperparameters, using larger batches in training, and removing the sentence pair prediction criterion, so it could focus on more to the point and wide-ranged masking during training. This in turn leads to a model that can better mimic spoken language, it is thus a model that can capture the fine intricacies of language making it the ideal choice for the task of sentiment analysis.

For sentiment analysis in connection with the personal diary relevance, moreover, RoBERTa provides advanced understanding of context and the capability to grasp fine nuances of linguistic expressions. This ability is critical in terms of identifying the sentiment of the diaries, especially when users may describe their emotions in various and complicated ways.

The most important reason that RoBERTa was chosen for mood assessment feature is just that it is extremely good at text processing being better in terms of precision. This allows the system to group text accurately into the most specific sentiment categories, identifying the message conveyed through written form to the deepest level of human feelings thus offering deeper insights into diary entries.

3) *Implementation Using Transformer Library and PyTorch*

The usage of the RoBERTa model for sentiment analysis in Oasis has the pleasure of relying on PyTorch and Transformer library, a powerful couple that can be used for NLP activities. The Transformer library offers ready-made models such as RoBERTa that developers can use with virtually no effort; as a result, project timelines are shortened, and complex aspects of NLP, such as model creation, are eliminated. PyTorch, which is used for building, training, and testing models, has become an important framework of choice due to its high flexibility and scalability.

The first step that is adopted by the developers to incorporate RoBERTa is using the Transformer framework to load the pre-trained model. PyTorch's adaptable computation diagram enables simple adjustment of the model to the traits of diary entries for sentiment analysis. This can be realized through the fine-tuning of a model specifically prepared for emotional text and it can be done by selecting dataset that contains very different sentiments expressed in student diary.

One the students have shared the experiences through the diary record, it forwards the text to the server application through a secure API. The server, on the back of PyTorch, with the RoBERTa model in place, evaluates the sentiment, and then the application receives the sentiment classification from the server. Firebase acts as a database that can save and retrieve information of the users' entries, and make sentiment analysis results available in real time, therefore users will enjoy getting feedback and recommendations within a short time interval as well as their emotional status.

C. High-Level Design of the Job Recommendation based on resume data

1) Resume Parsing Techniques and Job Recommendation Logic

- Resume Parsing Techniques: The parsing of resumes involves the extraction of critical information from resumes that are usually unstructured. Modern resume parsers use methods of NLP and ML to perceive required patterns, context, and entities in resumes. There are approaches like the Named Entity Recognition (NER) which provides the ability to categorize the pieces of text for predefined groups (for example, names, organizations, skills).
- Job Recommendation Logic: The processed information which is then structured is run through a job description matching algorithm to populate candidates' profiles with relevant job recommendations. Performing this task is narrowed down to looking through job adverts to obtain data consisting of major, skills, experience, and educational qualifications required by a specific job. The score provided by automated skills matching algorithms often matches the skills of the job to the candidate when using a similarity scoring algorithm that compares job requirements to the candidate's profile.

2) Implementation considerations

- Data Preprocessing: One of the key steps in the execution process is to carefully prepare representative data sets from resumes. This phase has the biggest impact regarding the transformation process: paper's textual information must be asked to take a form that can be understood by algorithms. Pre-processing is an important task, which includes among others the removal of irrelevant symbols and whitespace. Without this acting, the outcome of the research could be skewed. On top of that, tokenization divides the document into smaller parts, usually tokens, making it possible to have a low-level scrutiny. Part of speech tagging attributes grammatical functions to each token. Therefore, it provides a way to understand the text more deeply and precisely, since semantic extraction often depends on it.
- Training the NER Model: The key to the model is a NER (Named Entity Recognition) module, trained to be accurate after manually labeling specific entities present in datasets. This model is crucial for noticing what data could be extracted from the resumes and then providing that data. spaCy is very useful here in that NER stands out having an outstanding impact on an effective way of building custom models. Such efficiency comes from the fact that spaCy is optimized for real-world application deployment implying it becomes the most sought-after tool by developers who want to implement the state-of-the-art NLP with minimum resource efficiency.

D. High-Level Design of the Career Path Guidance

1) Understanding User Needs and Inputs

It is necessary to clearly identify the attributes needed for personalization. It is the data that has a lot of variances from one student to another. Identified attributes include the students major, academic interests and skills. It might seem quite ordinary, but each of the gathered statistics has a very important role when it comes to generating the individual career advice suggested by the system.

- Major/Course: Students' major or selected course reflects what they are naturally attracted to, revealing innate abilities and predispositions. Such data is helpful in building the momentum towards the identification of the possible industries and job roles that embody the students' preferences.
- Academic Interests: Considering student dreams and academic aspirations (preference for programming, technical writing, computer networking, etc.) is very important to identifying the best career path that is realistic and best suited.
- Skills: Students already acquired skills can be used to identify the best suited job while taking advantage of the existing skills to the maximum.

2) Overview of Vector Indexing

Vector indexing is one of the most advanced techniques used in computer science and data analytics for the management of data with high dimensions and at the same time, to accomplish the process of retrieving them quicker. On the one side, vector indexing means that elaborated data items are transformed into vectors (points in a multi-dimensional space) and on the other side, the process of multi-dimensional indexing is applied to them in the way that allows for searching and similarity evaluation. This turns out to be an especially important battle when one must tackle the loading and processing of big datasets for which typical search methods can be slow and prone to get stuck.

3) Technical Deep Dive into Text Similarity Algorithms

Text similarity tools are a fundamental part of natural language processing (NLP) as they are intended to determine, as opposed to compare, the degree to which two pieces of text relate to each other in terms of meaning or content. These algorithms undergo the process of analyzing the text and discover the feature set that consists of parameters like word frequency, syntax, and semantic meaning. From then, mathematical models like cosine similarity are employed to quantify the similarity between the feature sets.

The important aspect of matching user profile with their potential career path requires the use of text similarity algorithms. Every student profile has among other students major, academic interests and skills broken down into individual texts, while each career path has its own requirements in terms of job position, educational qualifications and experiences that are also broken down into separate text documents. The text similarity algorithms then go further to seek patterns, keywords, and semantic relationships which are used to compute the degree of congruence between the student profile and job requirement as opposed to mere keyword matching. With these, a student's abilities and dreams are comprehensively explored and tailored career paths are generated.

4) Role of Paraphrase-minilm-l6-v2

Paraphrase-minilm-l6-v2 model is a highly valuable resource as it is playing a significant role in this sphere. It scans and associates textual data in student profiles with the potential career requirements. By exploiting its sophisticated capability to recognize paraphrases and semantic equivalence, the model can realize potential matches between the student profile and career options that would otherwise be undetectable since its matching extends beyond keyword equivalence and recognizes deeper semantic two-way connections. It allows recommendation of a career based not only on a direct match between a student profile and job requirements but also on the general intent and meaning. Subsequently paraphrase-mini-l6m-v2 thus satisfies requirement of providing accurate and comprehensive as well as personalized career path guidance.

5) Implementation of Career Path Guidance Using Paraphrase-MiniLM-L6-V2

The process starts with the formatting of the input data to be fed into the system in a format that is suitable for further analysis. This involves standardizing the data from the student profiles and career descriptions. After providing the formatted input, the paraphrase-minilm-l6-v2 model is ready for processing. This step produces vector representations of both a job seeker's qualifications and the requirements of various career paths. These vectors grasp the semantic structure of the text and enable insight beyond pure surface-level matches based on word presence.

The system then leverages similarity metrics to match these vectors, discovering respective job descriptions which have the highest semantic similarity to the student profile. Even if the exact correspondences are not used in both texts, the semantic similarity between the descriptions is still displayed. This implementation is empowering the system to offer very comprehensive, accurate, and customized career guidance.

E. High-Level Design of the Academic Related Event recommendation

1) Understanding User Attributes

The key feature of the recommendation process involves understanding whether an available event is relevant to a student profile. Understanding the proper attributes to recommend the related events is crucial here. The attributes considered for this are student major, academic interests, and the distance to the event location.

2) Data Loading and Preprocessing with Pandas

Pandas, a robust python software for data manipulation and analysis, is used. The library's broad array of functionalities enables clean, thorough loading, as well as preparative processes, set up which is a prerequisite for a successful delivery of personalized learning and event recommendations.

Event details are extracted from different sources like databases or APIs that provide us with info on workshops, lectures, and conferences. Once student profile data and event data are fed into the systems, pandas are used to convert all these unstructured data to data frame format enabling advanced operations on the datasets with ease.

Data cleaning at this stage also starts with ending duplicates, managing missing values, and fixing gaps and errors in data. Panda provides a comprehensive functional support for the procedure of cleaning and verification which avoids data problems and makes the dataset more accurate. Additionally, data transforming using pandas is an effort to convert raw data into an understandable form that is beneficial for analysis. Such a process could create columns from existing data based on event classification based on attributes considered.

Data loading and preprocessing functionality of pandas not only serves to speed up the workflow but also greatly improves the system's recognition of user interests and connecting the users with similar events. A system through

which pandas are employed for this vital assignment is a system that promotes the quality of data that is fed directly into the recommendation algorithms so the correct and well-tailored event recommendations.

3) Supervised Classification Overview

The support system uses the supervised classification as a method to identify and provide individuals with the best study materials and events to attend. Supervised classification is an ML algorithm in which the learning principle is based on dealing with labeled training data and predicting the type of unlabeled data. With this approach, the algorithm is taught by consulting on the pre-create dataset in connection of input and corresponding outputs. During the training process, it gradually acquires an ability to map the input features onto the classes-determined outcomes, after which it can properly classify new and unseen data sources applying the acquired patterns.

4) Implementing Random Forest Classifier

The Random Forest (RF) classifier, which is an influential ensemble learning technique, is selected because of its unparalleled capabilities of dealing with complex datasets and producing accurate decisions. The foregoing classifier works by building several decision trees during the training phase and then selecting the most frequent class out of patterns classification by the mode of classes or regression by mean predictions. RF realizes lowering the risk for overfitting that occurs with single decision tree due to it distributing the multiple predictions, which in turn increase the reliability level of the recommended decisions.

RF classifier is based upon the vectors of the parameters retrieved from the student profiles and event data to identify patterns and correlations. The consideration of input variables can predict which new events are most favorable to each student. The advantage of the RF model is its ability to handle the anomaly and variations in data, and recommending the best options based on available data.

5) Integration with Sci-kit Learn for Model Development

The classifier is facilitated by sci-kit learn, a python library for machine learning with wide adoption. With a handy intuitive interface, model development is of course made available, from feature extraction and model instantiation to fitting the model with training set and using the model to make predictions on the new, untried data. Scikit-learn extensibility by connecting its routines with some python libraries (such as Pandas for data manipulation) makes it more powerful. Sci-kit Learn through the system makes the machine useful in implementing the RF classifier which gives reliant and adaptations recommendations for student to attend academic events.

F. Implementation Method

The development of Oasis is planned in Sprint fashion with Agile methodology. To incorporate real-time feedback and adjustments, the development process for this system is Incremental method. Therefore, the system is expected to cater to user needs satisfactorily. Development is done in stages adhering to the Modularity of system design, meaning the system is planned to be developed as separate components which then is integrated into a whole. The use of git (version control tool) made the programming efficient and a way through which the team could collaborate conveniently.

G. System Evaluation Method

The evaluation is planned in two stages as Functional Testing and Usability Testing.

1 Functional Testing: Verify that each functional unit performs as expected.

2 Usability Testing: This is the phase in which the system is tested in the real world by the intended audience. For this purpose, a System Usability Scale (SUS) test and a Focus-group interview are conducted.

2.1 SUS test: SUS is a free 10 item standardized usability questionnaire designed to measure users' perceived usability of a product or system. SUS score can tell systems' performance in the aspects of effectiveness, efficiency, and overall ease of use. The level of usability of the system could be identified based on the calculated final SUS score [31].

2.2 Focus-group interview: Focus-group interviews are performed as a planned discussion with a selected group of users and are conducted by a moderator. The participants for the interview are sampled from the study population. The aim is to obtain direct feedback from the users.

4. RESULTS AND DISCUSSION

Testing of the entire Comprehensive Undergraduate Support System – 'Oasis' is crucial in the realization that the final product is not only functional and reliable but also in accordance with the expectations and needs of the desired user group.

1. Functional Testing

The following sections provide results of the individual component accuracy and performance.

1.1. Proactive support for Mental Well-being

The bar plot in Figure 02 represents the distribution of samples across different sentiment classes in the dataset. The x-axis represents the sentiment classes, labeled as 0, 1, and 2, while the y-axis shows the frequency or number of samples for each class. The bar plot indicates that the dataset is relatively balanced, with each sentiment class containing a substantial number of samples. A balanced dataset is essential for training machine learning models, particularly in sentiment analysis tasks, as it helps prevent the model from becoming biased towards any class during the training process. If one class dominates the dataset, the model may learn to overfit that class and perform poorly on underrepresented classes.

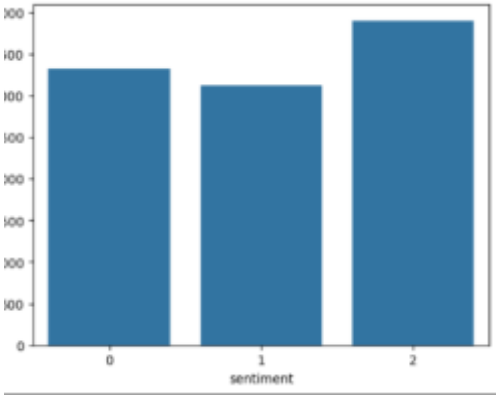

Figure 2. Sentiment distribution in dataset

Figure 03 presents a density plot illustrating the distribution of token counts (word counts) per sentence in the dataset. The plot showcases a right-skewed distribution, indicating that most sentences have a relatively small number of tokens, with a few outliers containing significantly more tokens.

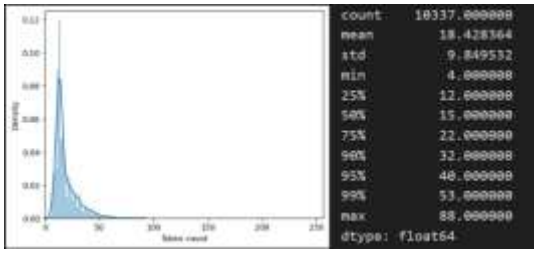

Figure 3. Illustrating the distribution of token counts.

The accompanying summary statistics provide further insights, revealing a mean token count of 18.4, a standard deviation of approximately 9.8, a minimum token count of 4, a maximum of 88 and the 50th percentile (median) of 15 tokens. This information is crucial for preprocessing steps like padding, as it helps determine a suitable maximum sequence length for model input when working with this dataset.

The confusion matrix presented in figure 04 provides a comprehensive evaluation of the performance of a sentiment classification model. The matrix compares the actual sentiment labels (fear, anger, and joy) against the predicted labels, allowing for a detailed analysis of the model's accuracy and potential areas for improvement.

The diagonal values represent the instances where the model correctly predicted the sentiment label. Notably, the model accurately classified 3303 instances of fear, 3107 instances of anger, and 3888 instances of joy. These high values along the diagonal indicate that the model has a strong ability to correctly identify the sentiment in most cases. The off-diagonal values represent misclassifications, where the model incorrectly predicted the sentiment label. For example, there were 16 instances of anger misclassified as fear, and 5 instances of joy misclassified as fear. However, the low values in these off-diagonal cells suggest that the model has a relatively low rate of misclassification errors.

	precision	recall	f1-score	support
fear	0.99	1.00	1.00	3314
anger	1.00	0.99	0.99	3125
joy	1.00	1.00	1.00	3898
accuracy			1.00	10337
macro avg	1.00	1.00	1.00	10337
weighted avg	1.00	1.00	1.00	10337

Figure 4. Confusion matrix.

The performance metrics presented in figure 05 further reinforce the model's effectiveness. The precision values (0.99 for fear, 1.00 for anger, and 1.00 for joy) indicate the proportion of positive predictions that were correct, while the recall values (1.00 for fear, 0.99 for anger, and 1.00 for joy) represent the proportion of actual positive instances that were correctly identified. The F1-score, which combines precision and recall into a single metric, is 1.00 for fear and joy, and 0.99 for anger, indicating an excellent balance between precision and recall for all

sentiment labels. The overall accuracy of the model is reported as 1.00, and the macro and weighted averages for precision, recall, and F1-score are also 1.00, further highlighting the model's strong performance across all sentiment classes.

These visualizations and metrics demonstrate the effectiveness of the sentiment classification model in accurately predicting sentiment labels with minimal misclassification errors. The high accuracy and low misclassification rates suggest that the model is robust and reliable in sentiment analysis tasks.

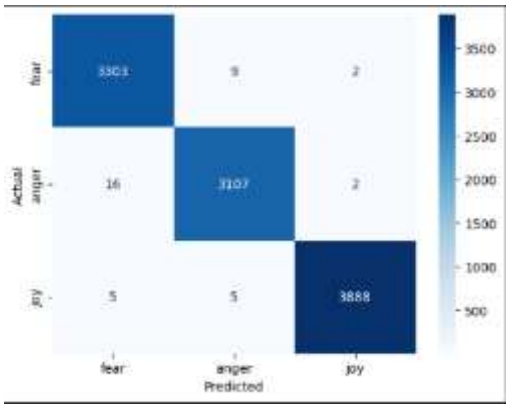

Figure 5. Performance metrics.

1.2. Job recommendations based on resume data.

The output in figure 06 provides a detailed view of the extracted information from a resume. After extracting relevant data from resume, the system recommends the most suitable jobs currently available in the job market to the user. These job vacancies are displayed on the user's home screen as given in figure 07, allowing them to view and consider potential employment opportunities.

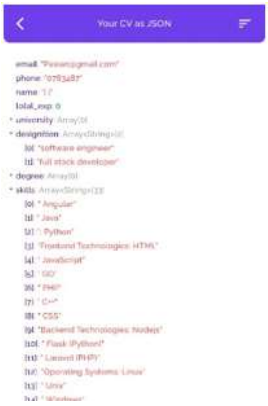

Figure 6. Extract text data from resume.

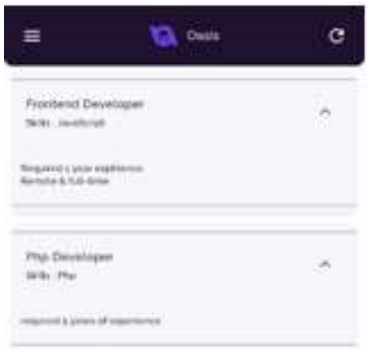

Figure 7. System recommends jobs.

1.3. Career path guidance

The dataset features such as student major and interests are prepared for the model. Each row of this data is analyzed and converted into a 384-dimensional vector. The resulting vector is depicted in figure 08. The top three highest similarity scores are then sorted and used to extract relevant career path guidance.

Figure 8. Dataset features.

This guidance as given in figure 09 is displayed to the user on a screen with the skills and qualifications needed to target that job.

Figure 9. Relevant career path guidance.

1.4. Networking through related event recommendation

The confusion matrix in figure 10 provides a visual representation of the performance of the classification model. In this matrix, the y-axis represents the true labels, while the x-axis represents the predicted labels. Each cell in the matrix shows the number of instances that are predicted as a certain class, with the true class being indicated by the row and the predicted class by the column. The values along the diagonal represent the instances that are correctly classified. From the given matrix, most of the values are concentrated along the diagonal, indicating that the model has high accuracy. Specifically, the majority of Data Scientists, Graphic Designers, HR Specialists, Marketing Managers, Software Engineers, and Web Developers have been correctly classified, with very few misclassifications. For example, out of 119 Data Scientists, 111 were correctly classified, with only 8 misclassified.

	precision	recall	f1-score	support
Data Scientist	1.00	0.93	0.97	119
Graphic Designer	1.00	1.00	1.00	115
HR Specialist	0.91	1.00	0.95	79
Marketing Manager	1.00	1.00	1.00	60
Software Engineer	1.00	1.00	1.00	118
Web Developer	1.00	1.00	1.00	83
accuracy			0.99	574
macro avg	0.98	0.99	0.99	574
weighted avg	0.99	0.99	0.99	574

Figure 10. Confusion matrix.

The classification report in figure 11 provides detailed metrics on the performance of the model for each class. It includes precision, recall, and F1-score for each category, along with the support values which represent the number of instances for each class. From the report, we see that the precision, recall, and F1-score are all very high, generally around 0.98 to 1.00, indicating excellent performance across all classes. For instance, the precision and recall for Graphic Designers are both 1.00, leading to an F1-score of 1.00. For Data Scientists, while the precision is 1.00, the recall is slightly lower at 0.93, resulting in an F1-score of 0.97.

The overall accuracy of the model is 0.99, which signifies that 99% of the predictions made by the model are correct. The macro average and weighted average for precision, recall, and F1-score are also very high, reinforcing the consistency and reliability of the model across different classes. The macro average treats all classes equally, while the weighted average considers the support (number of instances) of each class. Given the high values in both averages, it can be inferred that the model performs well.

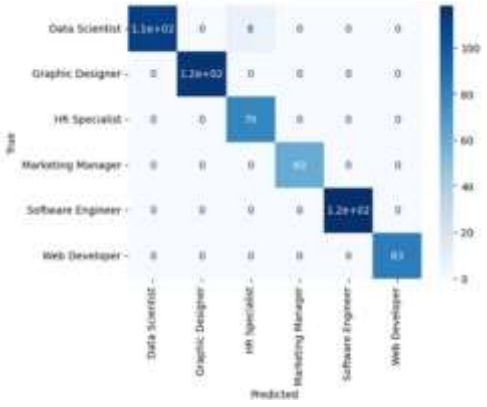

Figure 11. Performance metrics.

2. Usability Testing

2.1. SUS test

To determine the SUS score, participants answer 10 items in a standard SUS questionnaire designed to assess their perception of the system's usability. Users rate each item on a scale from Strongly Disagree to Strongly Agree, which allows for the calculation of a score for a given response. The scale points are as, Strongly Disagree - 1, Disagree - 2, Neutral - 3, Agree – 4 and Strongly Agree - 5.

To calculate the final score, the following steps need to be performed. An online SUS calculator is used to get the result up to Step 2.

- Step 1: Subtract the selected scale position from 1 on all oddly numbered items (item 1, 3, 5, 7, 9), and subtract 5 from the selected scale position on all evenly numbered items (item 2, 4, 6, 8, 10).
- Step 2: Add up the points from the 10 questions into a user's total points and multiply the user's total points by 2.5 to get the individual user's score.
- Step 3: Repeat steps 1-2 for all users, and then average all to get the final SUS score.

Based on the final SUS score obtained after step 3 above, and the standard SUS score Adjective Rating values as given in Table 1, the systems' level of usability is identified as 'Good'. This result justifies that the systems' usability is to an acceptable level.

Sub Score	Grade	Adjective Rating
>80.3	A	Excellent
68 – 80.3	B	Good
68	C	Okay
51 - 68	D	Poor
< 51	F	Awful

Table 1. General SUS interpretation guide.

2.2. The Focus-group interview with selected 4 undergraduate students from 4 different HEI, aims at getting more comprehensive feedback allowing them to interact with the live system. Students responded to the system with many positive comments, mentioning the uniqueness and variety of the support provided to students along with valuable insights for future versions. The following section highlights critical insights into the system's performance and areas for improvement based on the focus-group interview.

- Proactive support for mental well-being: Users expressed their positive opinion on providing students with a reliable platform to take care of their mental well-being effectively. They accepted that the system is capable of motivating students to a routine of daily reflection through interactive diary writing. Also, that the received feedback and personalized attention that is tailored to the individual's feelings is very supportive in building the students' self-esteem. But they had concerns about the privacy of the diary entries. Privacy & data security is of fundamental importance, and that is particularly relevant when disclosing sensitive information such as sharing a personal diary. The system should clearly define the usage policies on data, giving users peace of mind knowing that their personal data is being taken care of and treated with the greatest security and confidentiality measures by the system owners.

- Job Recommendation based on resume data: The users mentioned that this component helps them to easily identify the best matching job opportunities that they could apply for when needed for their internships. But certain other aspects that should be considered when recommending a job position or career path, like salary range, job benefits and location are highlighted. Also, two users identified certain mismatches with the job opportunities recommended to them with their qualifications. Parsing resumes is one of the major problems of resume parsing systems, as statements and resumes are filled with diversified and inconsistent format. The way information is presented is a formidable issue that is not easily resolved by automatic data extraction. Also catering to the dynamic requirements of the industry is also a concern. The increasingly fast-paced and ever-changing nature of the job market implies that the system should be able to extract and use timely data, which is a bigger challenge. Additionally, the comparison of the resume extracted data with job position requirements is an absolute necessity in predicting correct job recommendations. Such aspects speak loudly of an advanced machine learning pipeline and best of the data handling practices ensuring sustainability of the system with the time.

- Career Path Guidance, and Networking through related event recommendation: Users greatly appreciated the inclusion of these components and acknowledged this as very essential for undergraduate students. The most commendable feature of this system is mentioned as its capacity to recommend individual career paths that they would like to engage in future based on their major, academic interests and skills, along with the information on required skills and qualifications to target it. Also, users appreciated the related event recommendation components significant contribution towards the goal of motivating, broadening the students' horizons and paving them a path for long term success.

However, when considering the overall results, the system influence on users has been very much positive, and majority of the users have reported that the system is supportive for self-reflection of their mental status, identifying relevant career options, providing extra study support, and identifying collaborative opportunities, therefore being a great supporter to their undergraduate life releasing them of lot of hassle.

5. CONCLUSION AND FUTURE WORK

Oasis has emerged as a transformative solution for HEI in the landscape of Student Support Services, adeptly addressing the multifaceted needs of undergraduate students. Through its innovative integration of proactive support for mental well-being, job recommendation based on resume data, career path guidance and facilitation of networking through related event components, Oasis offers a holistic approach to student support that is both personalized and technologically advanced.

Using Oasis has resulted in excellent outcomes, with users discovering that the ability to have self-awareness on the mental status leading to early interventions, figuring out the current options in job placement and identifying other available career paths based on their major and interests, and identifying possible networking opportunities as very useful in the undergraduate journey. The system's application of the latest advanced technologies under ML and NLP has been instrumental in reaping this success, since these innovations supply students with personal and pertinent services that meet a variety of needs.

Even though Oasis is recommended as a system to cater to student support services necessary in HEI, it has its shortcomings too. The provided recommendations are solely based on the attributes considered in the datasets. The full potential of a student's capabilities, experiences and aspirations gets ignored due to this limitation. This could be reduced with the introduction of more relevant attributes for an enhanced personalized experience. In future, the system could be improved to provide more relevant job and academic event recommendations through integration of relevant online portals to get more reliable and UpToDate information on job adverts as well as event details. Also enlarging the degree of supported academic majors when identifying suitable jobs and academic related events could be considered. The accuracy of the recommendations could be further improved by installing feedback loops to get feedback on the acceptance/rejection of the recommendations provided by the system.

Based on the system evaluation results, the system has proved effective in addressing major areas where undergraduate students need support and its attempt to de-mystify the fear and uncertainties of moving from academic to professional world. Furthermore, the evaluation emphasized the need for iterative improvement of the system guided by user feedback and technological advancements.

5. ACKNOWLEDGEMENT

Our special thanks to all the lecturers, and instructors of the Sri Lanka Institute of Information Technology (SLIIT) for the knowledge so willingly shared, which have significantly inspired and enhanced our research work. Also, we are deeply grateful to our parents for their unwavering support and encouragement.

REFERENCES

[1] Lumadi, Rudzani. (2021). Enhancing student development through support services in an open distance learning institution: A case study in South Africa. South African Journal of Higher Education. 35. 10.20853/35-1-4422.

[2] Brown, K., & Jones, L. (2019). Meeting the Diverse Needs of Students: A Case Study of Student Support Services in a Large Urban University. International Journal of Student Affairs, 12(1), 78-91.

[3] Wang, Y., & Li, M. (2018). Exploring the Use of Online Student Support Services by Non-Traditional Students: A Case Study of Adult Learners in Distance Education. Distance Education, 39(4), 512-528.

[4] V. S. Salunke and K. J. Bazner, "Examining Effective Student Support Services for STEM Graduate International Student Success," 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE), Las Vegas, NV, USA, 2023, pp. 1163-1165, doi: 10.1109/CSCE60160.2023.00194.

[5] Taeza-Cruz, Maria Elisa Linda; Capili-Kummer, Marifel Grace. (2023). Decision Support System to Enhance Students' Employability using Data Mining Techniques for Higher Education Institutions. International Journal of Computing and Digital Systems. 10.12785/ijcds/1301102.

[6] Rotar, O. Online student support: a framework for embedding support interventions into the online learning cycle. RPTEL 17, 2 (2022). https://doi.org/10.1186/s41039-021-00178-4

[7] Labadze, L., Grigolia, M. & Machaidze, L. Role of AI chatbots in education: systematic literature review. Int J Educ Technol High Educ 20, 56 (2023). https://doi.org/10.1186/s41239-023-00426-1

[9] Garcia, M., & Martinez, L. (2016). Leveraging Technology for Student Success: A Framework for Implementing Online Student Support Services. Journal of Educational Technology Systems, 44(3), 289-306.

[10] Z. Min, "Big Data Driven Mental Health Assessment Model for College Students," 2022 International Conference on Knowledge Engineering and Communication Systems (ICKES), Chickballapur, India, 2022, pp. 1-5, doi: 10.1109/ICKECS56523.2022.10059983.

[11] W. Huang, C. Liang, M. Rola and X. Zhong, "Research on Mental Health of College Students Based on Branch and Bound Algorithm," 2023 4th International Conference on Information Science and Education (ICISE-IE), Zhanjiang, China, 2023, pp. 248-251, doi: 10.1109/ICISE-IE60962.2023.10456527.

[12] O. V. D. E. Roshani, H. M. K. J. J. Ayuwardhana, P. H. M. S. Rodrigo, R. U. Hewageegana, H. Fernando and D. I. de Silva, "Novel Approach for Enhancing Mental Well-Being Through Machine Learning Techniques," 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2023, pp. 651-656, doi: 10.1109/ICAC60630.2023.10417149.

[13] C. Senette, M. C. Buzzi and M. T. Paratore, "Self-assess Momentary Mood in Mobile Devices: a Case Study with Mature Female Participants," 2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 2023, pp. 1-7, doi: 10.1109/ICCT56969.2023.10075891.

[14] S. Hu, A. Kumar, F. Al-Turjman, S. Gupta, S. Seth, and Shubham, "Reviewer Credibility and Sentiment Analysis Based User Profile Modelling for Online Product Recommendation," *IEEE Access*, vol. 8, pp. 26172–26189, 2020, doi: 10.1109/ACCESS.2020.2971087.

[15] D. Bougharеb, R. Boughareb, N. Guerri, and H. Seridi, "FindLoc: A Sentiment Analysis-Based Recommender System," in *2023 3rd International Conference on Computing and Information Technology,*

ICCIT 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 101–105. doi: 10.1109/ICCIT58132.2023.10273931.
[16] Q. Ren, Y. Zheng, G. Guo, and Y. Hu, "Resource Recommendation Algorithm Based on Text Semantics and Sentiment Analysis," in *Proceedings - 3rd IEEE International Conference on Robotic Computing, IRC 2019*, Institute of Electrical and Electronics Engineers Inc., Mar. 2019, pp. 363–368. doi: 10.1109/IRC.2019.00065.
[17] H. Khatter, N. Goel, N. Gupta, and M. Gulati, "Movie Recommendation System using Cosine Similarity with Sentiment Analysis," in *Proceedings of the 3rd International Conference on Inventive Research in Computing Applications, ICIRCA 2021*, Institute of Electrical and Electronics Engineers Inc., Sep. 2021, pp. 597–603. doi: 10.1109/ICIRCA51532.2021.9544794.
[18] Y. Hongda and K. Takano, "A Recommendation Method for Social Media Users based on a Sentiment Analysis Model," in *LifeTech 2022 - 2022 IEEE 4th Global Conference on Life Sciences and Technologies*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 485–488. doi: 10.1109/LifeTech53646.2022.9754863.
[19] D. Ao and C. Zhang, "Deep Collaborative Filtering Recommendation Algorithm Based on Sentiment Analysis," in *Proceedings - 2023 8th International Conference on Information Systems Engineering, ICISE 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 403–408. doi: 10.1109/ICISE60366.2023.00092.
[20] N. Bali, K. Singh, and S. Dhawan, "A Survey and Comparative Analysis of Relevant Approaches of Recommendation System," Institute of Electrical and Electronics Engineers (IEEE), Jan. 2024, pp. 750–755. doi: 10.1109/ic3i59117.2023.10397781.
[21] K. Oooka and Y. Wang, "A Tourist Spot Recommendation Method Based on Image Search and Sentiment Analysis," in *GCCE 2022 - 2022 IEEE 11th Global Conference on Consumer Electronics*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 670–674. doi: 10.1109/GCCE56475.2022.10014382.
[22] L. Berkani, M. Zaouidi, and R. Brahimi, "Sentiment deep learning algorithm for multi-criteria recommendation," in *Proceedings of the 2022 1st International Conference on Big Data, IoT, Web Intelligence and Applications, BIWA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 77–82. doi: 10.1109/BIWA57631.2022.10037849.
[23] Mgarbi, Hanae; Chkouri, Mohamed Yassin; Tahiri, Abderrahim. (2023). Towards a New Job Offers Recommendation System Based on the Candidate Resume. International Journal of Computing and Digital Systems. 10.12785/ijcds/140103.
[24] Lakshan W.G.K; Prabuddhi K.A.V; Weerasinghe, P P; Bandara E.M.L.P; Gamage, M P; De Zoysa R.R.P. International Research Journal of Innovations in Engineering and Technology; Dharmapuri Vol. 7, Iss. 11, (Nov 2023): 307-314. DOI:10.47001/IRJIET/2023.711042.
[25] Tejaswini K, Umadevi V, Shashank M Kadiwal, Sanjay Revanna, Design and development of machine learning based resume ranking system,Global Transitions Proceedings, Volume 3, Issue 2, 2022, Pages 371-375, ISSN 2666-285X, https://doi.org/10.1016/j.gltp.2021.10.002.
[26] Girish K. Palshikar, Sachin Pawar, Anindita Sinha Banerjee, Rajiv Srivastava, Nitin Ramrakhiyani, Sangameshwar Patil, Devavrat Thosar, Jyoti Bhat, Ankita Jain, Swapnil Hingmire, Saheb Chaurasia, Payodhi Mandloi, Durgesh Chalavadi, RINX: A system for information and knowledge extraction from resumes, Data & Knowledge Engineering, Volume 147, 2023, 102202, ISSN 0169-023X, https://doi.org/10.1016/j.datak.2023.102202.
[27] Al Abri, Amal; Jamoussi, Yassine; AlKhanjari, Zuhoor; Kraiem, Naoufel. (2020). PerLCol: A Framework for Personalized e-Learning with Social Collaboration Support. International Journal of Computing and Digital Systems. 10.12785/ijcds/090312.
[28] P. Jeryluxshigan, J. Kunekanth, V. Shothripragash and Y. Sanojan, "BRIGHT MINDS Online E-Learning App for Kids - Web Application," 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2023, pp. 657-661, doi: 10.1109/ICAC60630.2023.10417282.
[29] K. Krishnananthan, A. Rasheed, R. Thadchaneswaramorthy, K. Sathiyavarathan, S. Thelijjagoda and K. Rajendran, "HANDTALK: Adaptive and Interactive Self-Learning System for Deaf and Dumb School Students in Sri Lanka," 2023 5th International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2023, pp. 739-744, doi: 10.1109/ICAC60630.2023.10417698.
[30] Olubunmi Ipinnaiye, Angelica Risquez, Exploring adaptive learning, learner-content interaction and student performance in undergraduate economics classes, Computers & Education, Volume 215, 2024, 105047, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2024.105047.
[31] John Brooke. 2013. SUS: a retrospective. Journal of Usability Studies 8, 2 (February 2013), 29–40.

Author 1 Sellapperumage Maneesha Fernando is an undergraduate student in the Computer Science and Software Engineering Department of the Sri Lanka Institute of Information Technology, specializing in Software Engineering. Outside of her academic pursuits, she enjoys experimenting with new software tools and taking long walks in nature.

Author 2 Paranagama Marankothlage Deshan Chathuranga Rajapaksha is an undergraduate student in the Computer Science and Software Engineering Department of the Sri Lanka Institute of Information Technology, specializing in Software Engineering. In his leisure time, he enjoys attending technology workshops and playing cricket with friends.

Author 3 Manori Gamage is an academic working in the Information Technology Department of Sri Lanka Institute of Information Technology. Her research areas are in educational technology, storage technologies and machine learning. She is passionate about exploring how the latest advancements in AI technology are transforming education and storage methods.

Author 4 Suduhakuruge Umesh Madushan is an undergraduate student in the Computer Science and Software Engineering Department of the Sri Lanka Institute of Information Technology, specializing in Software Engineering. During his free hours, he loves to engage in creative writing and attend tech meetups to

network with other professionals in the field.

Author 5 Kalagedara Dewage Lanka Jayamini Karunathilaka is an undergraduate student in the Computer Science and Software Engineering Department of Sri Lanka Institute of Information Technology, specializing in Software Engineering. In her free time, she enjoys reading articles on the latest advancements in technology and she likes to practice yoga.