Identifying Personality Traits using SVM Classifier and Resonance Neural Network

Umamaheswari J, Nivedha E, Santhosh A C
Dr. R. Vijayarajeswari
Department of Computer Science and Engineering
VELALAR COLLEGE OF ENGINEERING AND TECHNOLOGY, ERODE 638 012 (An Autonomous Institution Affiliated with Anna University, Chennai)

Abstract—the technical advancements of Data Science and AI have brought about a revolution across the globe. We can witness the application of machine learning in various aspects of our daily lives, and one of its most significant contributions is the ability to categorize individuals based on their personality traits. Every person has a distinct personality type, and the availability of vast amounts of high-dimensional data has enabled marketers to improve the effectiveness of their campaigns by targeting specific groups of people. Communication strategies based on personality types have proven highly effective in enhancing the appeal and prediction of products and services. In this project, we have developed a system that uses Big Five personality traits to predict personality types by analyzing human personality and personality features. Our proposed approach, which employs a hybrid support vector machine (SVM) and Resonance Neural Network (RNN), has been evaluated for accuracy against current state-of-the-art methods, both in laboratory and real-world settings, and has demonstrated greater accuracy.

I.INTRODUCTION

The NEO-PI-R or HEXACO-PI-R personality inventories, which are based on the five or six factor models of personality, are two frequently used methods for evaluating personality during the recruiting process. However, because they incur additional costs and are often less well-liked by candidates than other approaches, such as interviews, personality tests are not as frequently employed. Moreover, the interviewer's prejudice may be introduced during the interview. Structured interviews using standardized questions derived from job analysis have been found to reduce bias and enhance the capacity to predict work performance as a means of resolving this problem. Natural language processing techniques can be used to infer character from interview responses rather than depending on personality tests. While studies have shown that informal language use in online media can reveal personality traits, the consistency of these results in formal settings such as interviews has not yet been demonstrated.

II.EXISTING WORK

Existing Work Algorithm

A statistical technique called logistic regression is frequently employed in machine learning, especially for identifying activities in video sequences obtained from depth cameras. It is helpful for evaluating whether a certain activity is being performed or not based on local motion and shape information recovered from a depth sequence since it models the relationship between a collection of input variables and a binary output variable.

When applied to activity recognition, logistic regression can improve accuracy by combining neighboring excited sub personality formal's into a polynomial form. This enables the characterization of both local motion and shape information, enhancing the accuracy of activity recognition. However, it should be noted that images from depth cameras are not suitable for identifying unique individuals, but they can be used for activity recognition.

One of the main drawbacks of using depth cameras for activity recognition is the computational and time-intensive processing of large amounts of video data, making it impractical for real-time applications like surveillance or crime reduction. Although logistic regression can improve the accuracy of activity recognition, it may not be practical in such situations.

In spite of these restrictions, machine learning methods like logistic regression continue to demonstrate excellent promise in a range of applications, including image classification, natural language processing, and activity detection. These technologies are probably going to become more important as they develop in many different fields and sectors.

B.Disadvantages

This is primarily because each test statement provides a single data point for assessing a specific aspect of the underlying characteristic. Respondents become bored and irritated since it frequently demonstrates no

direct relation to the position an applicant is seeking for. It cannot be limited to self-rating on a predetermined set of facial expressions, as in typical personality tests.

C. Literature Survey

The authors of "A Meta-Analysis of the Links Between Personality and Workplace Deviance: Big Five vs HEXACO" did a meta-analysis of several research regarding the connection between personality and workplace deviation. The HEXACO Honesty-Humility domain was shown to be the best predictor of workplace deviance, followed by Conscientiousness and Agreeableness. They compared the accuracy of the Big Five and HEXACO personality models in predicting workplace deviance. HEXACO Emotionality and Big Five Neuroticism were also linked to workplace deviation. Overall, it seems that the HEXACO model could predict and explain levels of workplace deviation as well as the Big Five model.

The authors of "SMOTE TOMEK-Based Resampling for Personality Recognition" suggested a method for predicting personalities based on particle swarm optimization, synthetic minority oversampling, and Tome Link resampling (PSO-SMOTETomek). They discovered that this approach was effective with a limited sample, and personality detection accuracy was increased by up to 10%. The classification effects of short text data were shown to be superior to those of large text data. Current state-of-the-art models were greatly surpassed by the suggested model.

In the study "Personality Predictions Based on User Behavior on the Facebook Social Media Platform," the authors looked at the predictability of Facebook users' personality characteristics based on various aspects and metrics of the Big Five model. The connection between each of the feature sets and personality qualities was done after the analysis and comparison of four machine learning models. The results of the prediction accuracy test revealed that, even when evaluated using the same data set, the various machine learning models and feature sets varied in their ability to predict personality traits.

III. PROPOSED WORK

Proposed Work Algorithm

A deep learning-based approach for predicting human personality using image processing techniques. The accuracy of the approach is high due to the use of deep learning architectures, which work well with large datasets and can achieve low error rates. To improve the detection of personality, an efficient feature extraction method is used to generate a reliable background model. This also improves the time and speed of the process.

In order to increase the likelihood of successfully identifying personality, the method includes producing a huge number of frames to compare. The deep learning-based Kalman and Gaussian filters are used to carry out the preprocessing operations. The Histogram of Gradient (HOG) technique is used to extract features. Support Vector Machines (SVM) and Resonance Neural Networks (RNN) are used in a hybrid machine learning technique for classification.

Personality prediction is a crucial application of image processing, particularly in the field of security. The advantage of this algorithm is that it can be applied to any images, even if the input is partial. Overall, this approach offers a powerful tool for accurately predicting human personality from images using advanced deep learning techniques and sophisticated image processing algorithms.

S.No.	Logistic Regression (Existing)	Support Vector Machine (Proposed)
1	It is an algorithm that is used to address classification related issues.	It is a method for dealing with regression and classification problems.
2	It is supported by a statistical methodology.	It is based on the data's geometrical characteristics.

3	It functions with known independent variables.	Text and image-based semi- and unstructured data are both well suited for it.

Table 1: Difference between Logistic regression and Support Vector Machine

Advantages

One of the purposes of using validation techniques is to ensure the precision of the models that have been trained. Lower correlations in studies using image data could be due to a high degree of vocabulary variation. The accuracy achieved in predicting personality through image analysis is noteworthy and highlights the potential of this approach. Research has shown that interviews are more preferred by candidates than traditional personality tests. By extracting personality traits from interview responses, the need for a separate personality test can be eliminated, resulting in improved candidate satisfaction and engagement.

Modules and Description

Image Acquisition and Preprocessing

Image acquisition is the process of capturing an image using a camera or other imaging device. In many cases, this involves controlling various parameters of the imaging system, such as exposure time, aperture size, and gain, to capture an image with optimal contrast, brightness, and sharpness. Additionally, the selection of an appropriate lens and positioning of the camera or imaging device can also play a critical role in capturing high-quality images.

Once an image has been acquired, it may need to be preprocessed before it can be used for further analysis. The specific preprocessing techniques used will depend on the application and the characteristics of the image being analyzed. For example, in medical imaging, preprocessing may involve removing artifacts such as motion blur or patient movement to ensure accurate diagnosis and treatment planning. In contrast, in surveillance or security applications, preprocessing may involve segmenting the image into regions of interest and enhancing the contrast of these regions to better detect and track objects.

Image smoothing is a common preprocessing technique that involves applying a filter to the image to remove noise and make it more uniform. This can be particularly useful in applications such as medical imaging or remote sensing, where noise can interfere with the detection of critical features. Similarly, image enhancement techniques can be used to adjust the contrast and brightness of the image to make it easier to analyze, or to highlight specific features of interest.

Image segmentation is another important preprocessing step that involves dividing the image into regions that correspond to different objects or regions of interest. This can be particularly useful in applications such as object recognition or tracking, where it is necessary to isolate specific objects or regions in the image for further analysis. Feature extraction is another critical step that involves identifying and extracting features from the image that are relevant to the task at hand, such as edges, corners, or texture patterns.

In computer vision and image processing, image acquisition and preprocessing play crucial roles because they allow researchers and practitioners to draw out important information from images and videos that can be used to inform a variety of applications in industries like security, transportation, and healthcare. It is feasible to increase the precision and efficiency of the subsequent analysis phases and open up new possibilities for innovation and discovery by carefully managing the picture capture process and utilizing the proper preprocessing techniques.

Histogram of Gradient

Using HOG, preprocessing is carried out. The feature descriptor known as HOG, or Histogram of Oriented Gradients, is often used in computer vision and image processing for applications like object detection and recognition. Each bigger picture block's histogram of these orientations is formed by computing the gradient orientation of tiny image patches. These histograms are then normalized and combined to create the final feature vector for the whole image, which can then be used as input to a machine learning algorithm for object detection and recognition. HOG is particularly effective in detecting objects with unique textures or patterns, such as human bodies or faces, and is used in various applications such as robotics, self-driving cars, and surveillance. Although HOG is computationally efficient and straightforward compared to other feature descriptors, it may not be as effective in identifying objects with less distinctive patterns or in environments with significant variations in lighting or other environmental conditions. Therefore, it is

common to combine HOG with other feature descriptors or preprocessing techniques to enhance its accuracy and robustness in various applications.

II. Image Segmentation

Image segmentation, a fundamental stage in image analysis, groups pixels into meaningful sections based on their shared properties, simplifying an image. A variety of subsequent analyses, including object recognition, feature extraction, and classification, may be made possible as a result.

One of the main challenges in image segmentation is selecting an appropriate method for a given image and application. Each method has its strengths and weaknesses, and the choice of method can have a significant impact on the quality and accuracy of the segmentation results. For example, thresholding can be effective for simple images with distinct foreground and background regions, but it may not work well for more complex images with overlapping regions or varying illumination conditions. On the other hand, systems based on machine learning, such CNNs, can achieve great accuracy but may need a lot of computer power and training data.

Another challenge in image segmentation is dealing with noise, occlusions, and other factors that can affect the appearance of the image. Preprocessing techniques such as image smoothing, enhancement, and normalization can help mitigate these effects and improve the accuracy and robustness of the segmentation results.

Image segmentation has a broad range of applications in various fields, including healthcare, security, entertainment, and manufacturing. For example, in healthcare, segmentation is used to identify and quantify abnormalities in medical images, such as tumors or lesions. In security and surveillance, segmentation is used to detect and track objects of interest in real-time, such as vehicles or people. In entertainment, segmentation is used to create special effects and animations, such as green-screen compositing. In manufacturing, segmentation is used to inspect and analyze products, such as electronic components or food items.

Overall, image segmentation is a fundamental task in computer vision that enables researchers and practitioners to extract useful information from images and videos that can inform a wide range of applications.

K-Means Clustering

K-Means Clustering Algorithm is used for image segmentation. Unsupervised machine learning algorithms like K-Means Clustering, which divides comparable data points into clusters based on similarity, are often utilized. The method looks for the best K clusters, where K is chosen by the user, to reduce the sum of squares inside each cluster.

The K-Means approach first randomly initializes K centroids, one for each cluster, in order to do this. The next step is to allocate each data point to the closest centroid depending on how far away it is from the centroid. Until every data point has been allocated to a cluster, this procedure is repeated.

The approach computes the mean of all the data points assigned to a cluster after the first assignment of data points to clusters, updating the centroids. Until convergence, which is defined as no further change in the assignment of data points to clusters, the assignment and centroid update phases are repeated.

The effectiveness of the K-Means algorithm depends critically on the selection of K. A small K might result in oversimplified clusters, whereas a big K could lead to over fitting and the production of an excessive number of clusters. As a result, it's crucial to apply a suitable measure, such as the elbow method or the silhouette score, to determine the ideal amount of clusters.

The K-Means technique is more sensitive to outliers and unimportant data properties. Consequently, before using the K-Means technique, it is crucial to preprocess the data to eliminate noise and unimportant characteristics. To decrease the amount of features in the data, dimensionality reduction techniques like Principal Component Analysis (PCA) can also be utilized.

Marketing, image processing, and anomaly detection are just a few of the industries where K-Means clustering has a wide range of applications. K-Means clustering, for instance, may be used in marketing to categorize clients based on their demographics or purchase habits. K-Means clustering can be used in image processing to divide similar pixels into distinct regions. K-Means clustering may be used in anomaly detection to find data points that differ considerably from the other data points.

Overall, the K-Means technique is a strong clustering analysis tool that has been successfully used in a number of applications. To guarantee the algorithm performs at its best, it is crucial to carefully select the number of clusters, prepare the data, and assess the outcomes.

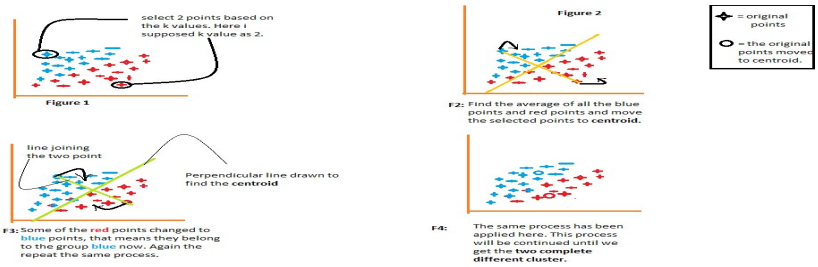

Figure 1: K-Means Clustering

III. Image Classification and Personality Prediction

In recent years, researchers in the domains of psychology, neuroscience, and artificial intelligence have paid increasing attention to the categorization and prediction of personality traits using machine learning algorithms. Large data sets may be analyses by machine learning algorithms, which can also spot patterns and connections that would seem obscure to people. Researchers can better grasp how various elements contribute to a person's personality by training these algorithms on data on personality traits.

Self-report questionnaires, peer ratings, and behavioral observation are three commonly used methods for collecting data related to personality traits. Self-report questionnaires are easy to administer and can provide a comprehensive assessment of an individual's personality. Peer ratings can provide a more objective assessment of an individual's personality, as they are based on the perceptions of others. Behavioral observation can provide valuable insights into how an individual's personality traits manifest in real-world situations.

Machine learning algorithms can be used to analyze data collected using these methods and identify patterns related to an individual's personality traits. For example, SVMs can be used to classify individuals into different personality types based on their responses to self-report questionnaires. Random forests can be used to identify the most important features related to personality traits, while neural networks can be used to make predictions about an individual's personality based on their behavioral observations.

Several sectors can benefit from the classification and prediction of personality characteristics. Companies may utilize personality features in marketing and advertising to specifically target certain customer groups with their advertising efforts. In human resource management, personality traits can be used to evaluate job candidates' fit with the company culture and to identify potential areas of conflict. However, it is important to note that machine learning algorithms are not a replacement for human judgment and should be used in conjunction with other methods to gain a more comprehensive understanding of an individual's personality traits.

A) FOR CLASSIFICATION: SUPPORT VECTOR MACHINE

Here, a support vector machine is used for categorization. A well-liked supervised learning technique for classification and regression applications is called a Support Vector Machine (SVM). The optimum hyper plane that divides data points into distinct classes with the widest possible gap between the classes is what SVM seeks to identify.

In order for SVMs to function, the input data points must be mapped into a high-dimensional feature space, and the hyper plane that maximizes the margin between the two classes must then be found. Support vectors, which are utilized to determine the decision border between the classes, are the data points closest to the hyper plane.

By utilizing kernel functions to shift the input into a higher dimensional space where a linear decision boundary may be discovered, SVMs are able to handle both linear and non-linear classification tasks. The performance of the SVM can be considerably impacted by the kernel function selection.

Text classification, picture classification, and bioinformatics are just a few examples of the many uses for SVMs. When working with high-dimensional data and few training examples, they are especially helpful.

SVMs are a potent technique for classification and regression problems in general, and several real-world applications have shown how successful they are. To attain the best performance, they might be computationally demanding and need careful hyper parameter selection.

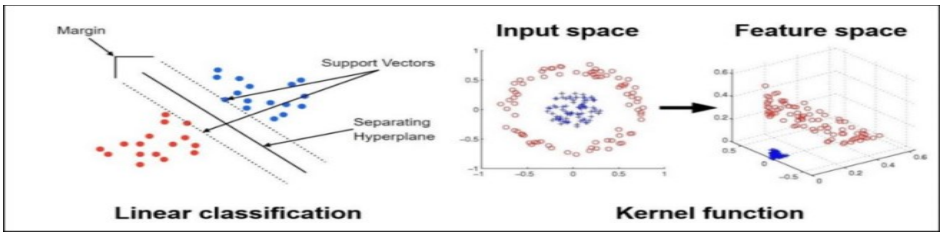

Figure 2: Support Vector Machine

The SVM algorithm is a popular machine learning technique used for personality prediction. It involves finding the optimal weight vector and bias term that can classify the input data points into different classes, using a decision boundary with the maximum possible margin between the classes. This is done by solving an optimization problem that aims to minimize the classification error and maximize the margin between the decision boundary and the closest data points of each class. Once the SVM classifier is trained on a dataset, it can be used to predict the personality traits of new individuals based on their input data, which usually consists of a set of features extracted from personality questionnaires or other sources of data. Based on the decision boundary learnt from the training data, the classifier utilizes these features to predict the personality traits of the individual.

$$y = \text{sign}\,(w * x + b)$$

Where:

- For the input data point, y is the expected output or class label (-1 or 1).
- The relevance of each feature in the input data point is determined by the weight vector w.
- The supplied data point, x, is shown as a feature. The bias term or intercept for vector b

B) FOR PREDICTION: RESONANCE NEURAL NETWORK

RNNs are based on the concept of resonance, which refers to the tendency of a system to vibrate with greater amplitude in response to certain frequencies. In the context of neural networks, resonance can be thought of as a form of selective attention, where the network focuses on specific patterns in the input and amplifies them.

RNNs consist of a network of interconnected neurons, where each neuron is associated with a specific input pattern. The neurons are organized into groups called resonance cells, where each resonance cell corresponds to a particular input pattern. When an input pattern is presented to the network, the resonance cells that match the pattern become active, and the network as a whole resonates with the input.

One of the advantages of RNNs is that they are particularly good at recognizing and classifying complex patterns with multiple layers of structure. For example, in speech recognition, RNNs can be used to identify phonemes (the individual units of sound in a language) and combine them into words and sentences.

Another advantage of RNNs is their ability to learn and adapt to new patterns over time. This is because the resonance cells can be trained to recognize new patterns and adjust their responses accordingly.

However, RNNs can be computationally expensive, especially for large datasets. They also require a lot of training data to perform well. Additionally, because RNNs are often used for tasks that involve complex patterns, they may not be as interpretable as some other types of neural networks, making it difficult to understand how they make their decisions.

Overall, RNNs are a powerful type of neural network that have been applied successfully in a variety of fields. As with any machine learning technique, their performance depends on the specific task and the quality and quantity of the data available.

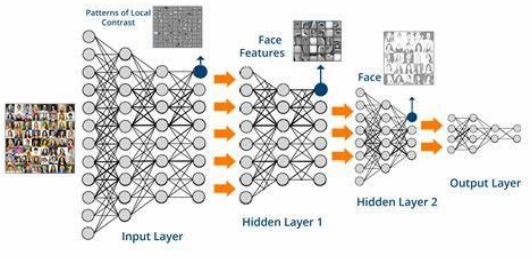

Figure 3: Resonance Neural Network

RNNs, a type of artificial neural network, are a suitable tool for personality prediction tasks. The formula for RNN can be represented as

$$h\_t = f(W\_x * x\_t + W\_h * h\_t\text{-}1 + b)$$

where the weight matrices W x and W h, which gauge the relative relevance of the input and the preceding hidden state, respectively, are used to compute the hidden state at a given time point. The model processes the input data sequentially, capturing the temporal dependencies and suitable for analyzing time series data such as personality trait trajectories over time. Once trained, the RNN model can predict the personality traits of new individuals based on their input data, which is typically a sequence of features extracted from their responses to personality questionnaires or other data sources, and then use the learned temporal patterns to predict the individual's personality traits.

IV. ARCHITECTURE

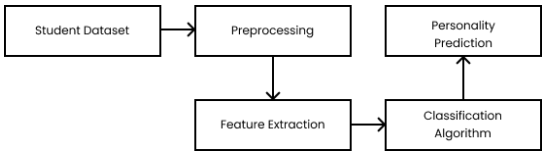

Figure 4: Architecture of Personality Prediction

V. SYSTEM REQUIREMENTS

A. Hardware Specifications

Processor Type : Core i5

RAM : 8 GB

Hard Disk : 500 GB

Keyboard : Standard Keys

Mouse : Optimal Mouse

B. Software Specifications

Operating System: Windows Family

Software : MATLAB

Database : SQLite

Dataset : Cagle

VI. APPLICATIONS

Here are some additional elaborations on the applications of personality prediction:

Healthcare:

The identification of patients who may be at danger of not adhering to their treatment regimens can also be aided by personality prediction, according to healthcare specialists. For instance, those with low conscientiousness levels may be less likely to adhere to their treatment plans. Moreover, personality analysis can assist medical practitioners in better understanding the preferences of their patients for particular types of therapies, resulting in more individualized and efficient care.

Financial services:

The Personality prediction can also be used to help financial institutions identify fraud and manage risk. For example, people who score high on openness to experience may be more likely to take risks, while those who score high on neuroticism may be more prone to financial anxiety and risk aversion. By understanding these personality traits, financial institutions can develop better risk management strategies.

Social media:

The identification and management of dangerous content on social media sites can also be assisted by personality prediction. For instance, those with high aggressiveness scores may be more inclined to harass others online, whereas those with high agreeableness scores may be more likely to connect with others in a good way. Social media companies may take action to restrict harmful information and foster good relationships by detecting certain personality traits.

Sports:

Sports organizations may improve their teaching and training methods by using personality prediction. For instance, those with high conscientiousness ratings may be more likely to stick to their training schedules, whereas individuals with high neuroticism ratings may be more prone to worry and stress. Sports teams may better satisfy the demands of individual athletes by adjusting their training and coaching by recognizing these personality qualities.

Customer service:

Moreover, personality prediction may be utilized by organizations to comprehend and better serve their clientele. For instance, those with strong extraversion scores may choose sociable and engaging customer service encounters, whereas individuals with high neuroticism scores may favor low-stress and effective interactions. Businesses may create better customer service strategies that cater to various consumer segments by studying these personality attributes.

In conclusion, personality prediction offers a wide variety of possible applications. We can leverage the potential of personality prediction to significantly enhance people's lives if we use these technologies appropriately and ethically.

VII. CONCLUSION

Personality prediction models have a broad range of applications, including in competitive exams and the military. They have the potential to provide individualized support and guidance to candidates by pinpointing areas that require improvement and offering focused training opportunities. Furthermore, personality prediction models can enhance the fairness and effectiveness of exams by determining the personality traits that are critical for success and developing tests that align with these characteristics. Ultimately, by identifying candidates with essential personality traits such as leadership, discipline, and decision-making abilities, these models can enhance the quality of recruits and boost the effectiveness of the armed forces. As additional data is gathered and analyzed, the precision of personality prediction models is expected to improve, enhancing their usefulness in guiding individuals towards their desired career paths.

REFERENCES

[1]. Borkenau, P., Brecke, S., Möttig, C. & Paelecke, M. Extraversion is accurately perceived after a 50-ms exposure to a face. J. Res. Pers. 43, pp.703–706(2009).https://www.academia.edu/27488602/Extraversion_is_accurately_perceived_after_a_50_ms_exposure_to_a_face

[2]. G. Matthews, I. Deary, and M. Whiteman, Personality Traits. Cambridge University Press,2009.https://www.cambridge.org/core/books/personality-traits/E78160A9C99B2B0452FF52D81D61C7B4

[3]. Kramer, R. S. S., King, J. E. & Ward, R. Identifying personality from the static, non expressive face in humans and chimpanzees: Evidence of a shared system for signaling personality. Evol. Hum. Behav (2011).https://psycnet.apa.org/record/2011-08073-004

[4]. Walker, M. & Vetter, T. Changing the personality of a face: Perceived big two and big five personality factors modeled in real photographs. J. Pers. Soc. Psychol. 110, pp.609–624(2016).https://www.researchgate.net/publication/281618676_Changing_the_Personality_of_a_Face_Perceived_Big_Two_and_Big_Five_Personality_Factors_Modeled_in_Real_Photographs

[5]. Naumann, L. P., Vazire, S., Rentfrow, P. J. & Gosling, S. D.Personality Judgments Based on Physical Appearance. Personal. Soc. Psychol. Bull. 35, pp.1661–1671(2009).https://journals.sagepub.com/doi/10.1177/0146167209346309

[6]. Ameta-analysis of the relations between personality and workplace deviance: Big Five VersusHEXACOhttps://www.sciencedirect.com/science/article/abs/pii/S0001879119300582

[7]. All the Dark triad and some of the Big Five traits are visible in the face

[8]. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.

[9]. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.

[10]. Personality predictions based on user behavior on the facebook social media platform https://ieeexplore.ieee.org/document/8494744

[11]. Personality recognition on social media with label distribution learning https://ieeexplore.ieee.org/document/7956172

[12]. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.

[13]. Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis and Implementation of 89C51 Controller Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools & Technology. 2022; 12(2): 34–41p.

[14]. C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.

[15]. G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756

[16]. G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:744-749