Personality Recognition & Video Interview Analysis

Supriya Anand
SIT (Department of CSE)
Symbiosis International University,
Pune- 411045

Nihar Gupta
SIT (Department of CSE)
Symbiosis International University,
Pune- 411045

Mayesh Mulay
SIT (Department of CSE)
Symbiosis International University,
Pune- 411045

Abhimanyu Sherawat
SIT (Department of CSE)
Symbiosis International University, Pune- 411045

Asst. Prof. Rupali Gangarde
SIT (Department of CSE)
Symbiosis International University, Pune- 411045

Abstract— **With artificial intelligence (AI) system we can predict human skill and its recommended job type as currently industry with employment agency needs to have AI introduced in it. With the help of all machine learning algorithms such as Naïve bayes, random forest, SVM, of convolutional neural network (CNN). Emotion of an employee, employees voice converted to text, face based company verification process will be done in this project. In the future system user can access user record ana analyzed the efficiency and how effective it is. Students also can get benefit by analyzing themselves with emotion and personality with this system as in the future we might see such system in the interview process. In the future system user can access user record ana analyzed the efficiency and how effective it is. Students also can get benefit by analyzing themselves with emotion and personality with this system as in the future we might see such system in the interview process.**

Keywords — Convolutional Neural Network (CNN), Company verification, Emotion Recognition, TensorFlow Naïve bayes, Random Forest, SVM,

INTRODUCTION

Interview is lengthy process nowadays. Employees going at place and giving interview is hectic in pandemic no proof maintained by any company with the interviewer like how he is answering, what he answered. Artificial intelligence system helps us in many ways to make things easy. In the traditional method of interviewing a candidate for certain job, the HR department of an organization invites a candidate Based on their resume for recruitment. This HR department manually analyzes the skills of candidates through their resume and see if he/she is fit for the required job. They conduct interviews and the panel has an important responsibility in assessment of the right candidate which would be fit for the job. In interview they not only check the skills, but also personality of the candidate as all the candidates recruited must have the right attitude and discipline for the Job. Developing a system which can identify person face emotion, voice emotion and analyzing resume are primary task of many similar this

application. So, we have decided to build an end to end interviewing the system with the help of AI which would not only classify resume of candidates but also recognize certain personality traits of candidates through asynchronous video interviews [1].

Although work completed before this time are related to many machine learning algorithms which are very time-consuming and might affect the system performance. CNN has proven great performance in an image processing task. CNN's model to create an automated AI-based interviewer that can automatically classify the applicant. The proposed work classifies the performance of a person based on video analysis in an interview. Finding the various features of facial expression, resume analysis and parsing, tone analysis speech emotion recognition and displaying it on user's system [2].
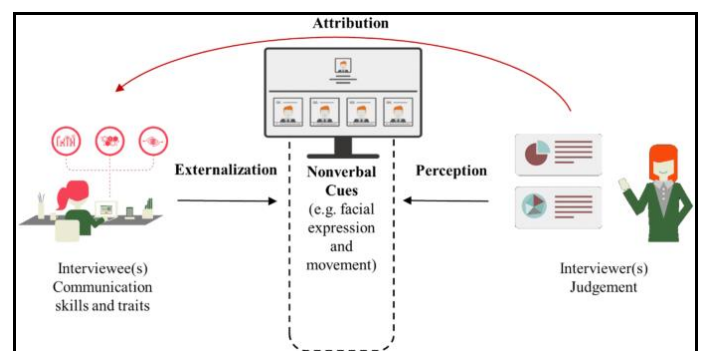

Fig 1: The process of interviewers 'judgments toward interviewees' communication skills and traits

LITERATURE SURVEY

In the previous research, Multimodal First Impression Analysis with Deep Residual Network author Yagura G, Isabelle Guyon, acknowledged a number of models for regressing the sensory data and the language data to the trait annotations and the interview annotations. The models that were used for predicting trait and interview annotations from the data in the same modality shared the same architectures

except for their penultimate layers. They presented several models for predicting the personality using short YouTube videos.

By authors Hxsung-Yufe Suhen, Kuho-En Hugng and Chimen-Liang Lin, in their paper Intelligent video interview agent used to predict communication skill set and personality traits, they developed an asynchronous video mode interview with AI based on Tensor Flow convolutional neural network (CNN), called as AVI-AI, that can be used to displace human raters.

In another paper authors Harari, Ramona Schoedel, Sumer Void, Samuel D. Gosling provided a brief overview of past studies related to personality recognition and use of it in job interviews. They provided an overview of studies did in past using machine learning in the use of job interview or learning the personality psychology. Then they acknowledged main challenges researchers have to face when interpreting, building and validating machine learning models.

Authors Dan Saadat, Butuan Balti and Dan Shiferaw used machine learning algorithms to identify raw input of image to text. They used CNN algorithm with various method to train a model that is able to accurately identify words. In the paper Who Am I? Personality Texts they identify the personality character traits from online text written by themselves. Personality detection based on texts from online networks has attracted many attentions.

The impact of AI within the recruitment industry: Defining a new way of recruiting by Dr David Atkinson and James Frisket, as the recruitment industry have significant issue the traditional recruitment processes are used to be found ineffective.

BACKGROUND

3.1 Interpreting CNN Model

Convolutional Neural Networks (CNN) are a form of deep artificial neural network that is very common. When compared to other image processing algorithms, CNN's use less preprocessing. The CNN's connectivity pattern parallels that of an animal's visual cortex.

Key components of a CNN: A convolution methodology that divides the precise analytical characteristics of the input. A fully connected layer that uses convolution layer performance to predict best definition, CNN design is galvanized by cortical area practicality and organization and designed to mimic vegetative cell property patterns.

Within a CNN the neuron area unit is divided into a 3D structure. That assortment of neurons analyzes a picture attribute, or every category of neurons focuses on recognizing one a part of the image. The CNNs then use the layer-based predictions to come up with a final output. It provides a vector of probability scores to mirror that a specific attribute belongs to a specific category.

The neural network CNN has AN design influenced by the cortical area of primates. The cortical area has many levels of the cortex (layers), every able to acknowledge additional

organized info. The specificity of the CNN is that the existence at all-time low of the stack (sequence) of the neural network layers of the coevolutionary (two-dimensional) layer followed by the pooling layers, as CNN combine layer.

3.2 Interpreting DNN Model

An interpretation is the transformation of an abstract definition (for example, a projected class) into a domain that humans can understand.

Deep neural network is a chain of neurons arranged in different layers, [3] with each layer receiving the previous layer's neuron activations as input and performing a simple computation. The network's neurons work together to create a dynamic nonlinear mapping from input to output.
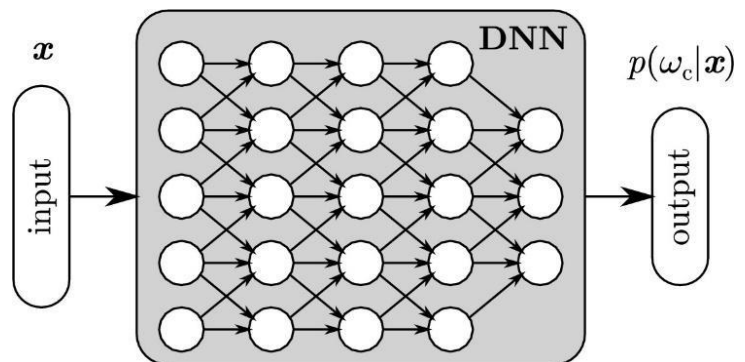

Fig 2: Example: neural network composed of the many interconnected neurons, which has the input x.

3.3 Resume Parsing

Resume parsing is a technology that helps us to intelligently extract data from online resumes. It enables recruiters to handle electronic resumes sent over the internet more effectively. It makes the process of resume and application screening process much simpler. Contact information, related skills, job history, and educational background can all be extracted with the aid of resume parsing technology. With up to 95% effectiveness, some resume parsing programmed achieve "near human accuracy.

"A resume is often thought of as a group of data associated with — expertise, instructional background, skills, and private details of an individual. Rules-based parsers don't stand an opportunity and an intelligent algorithmic rule is needed to extract text from raw documents. Optical Character Recognition (OCR) in conjunction with Deep IP algorithms on high will facilitate extracting the specified text. [4] Maintaining the vocabulary employed in resumes may be a huge challenge.
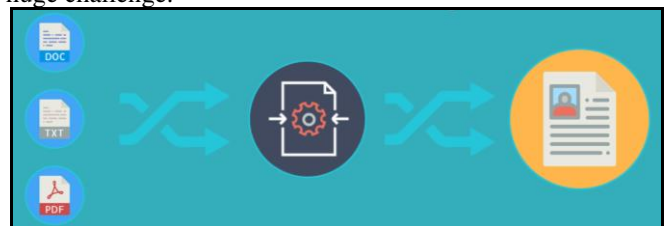

Fig 3: Resume Parsing Process

A resume consists of company names, establishments, degrees, etc. which might be written in many ways. For e.g., Skill: Skillate.com — [5] each of these words talk to an equivalent company, but are going to be treated as completely different words by a machine.

Deep data Extraction is an algorithmic rule that deep learning is often applied to, for data extraction in resumes. The basis of the matter here is knowing the context of a word. Applying Deep Learning to resolve data extracted greatly helped to effectively model the context of each word in an exceedingly resume.

3.4 Speech Emotion Recognition

In everyday interpersonal human relations, emotion plays a significant role. This is essential for both logical and intelligent decisions. [6] By sharing our feelings and offering input to others, it helps us match and appreciate the feelings of others. Emotion plays a major role in influencing human social interaction, according to research. Emotional displays offer a wealth of knowledge about a person's emotional state. This has spawned a new field of study known as automatic emotion recognition, whose primary aim is to comprehend and retrieve desired emotions.

There's now a Hand-Crafted Speech Emotion Recognition (SER) based feature where several researchers worked on handcrafted applications that used SER. Some researchers have techniques developed to identify speech emotion as data using spectrograms. SER has also been used by the Convolutional Neural Network (CNN) and CNN Alex Net models. It uses a complex model, the DBN used in speech learning applications, and the SVM classifier to define the emotional condition of the person. Research has suggested that CNN model can be used to predict the discrimination function from the entire utterance and can be subsequently used in an LSTM method to sequentially deduce emotions.

The primary purpose of the study is to perform a psychological survey ranging from psychology to computing. The data is being trained to recognize different emotions based on voice tone. [4] Two approaches for determining a person's different personality characteristics are self-rating and observer-rating. Organizational attitudes and results are projected using self-ratings in the psychology literature.

3.5 Facial Emotion Recognition

Facial recognition is a biometrics-based technology that verifies human faces by mapping facial features from a picture or video. [6] It compares the details and important information to a database of selected faces to find a suitable match.

A photo, video, or a good camera is used to take an image of the face. Facial recognition software creates a diagram of the face's geometry. [7] The distance between facial features is measured with this mapping. This recognizes facial landmarks that identify the face and allows for the development of a facial signature. Faceprints are linked to images in a database of facial authentication systems. A camera searches for the distinguishing features of a face (two eyes, a nose, and a mouth), and algorithms assist in determining the orientation and movement of the face.

Identification of a person's face. These apps operate by storing a picture of a person's face and then measuring the unique features that will enable it to be identified.

In pattern recognition and classification, machine learning algorithms have proved to be extremely useful. The features are the most critical aspects of any machine learning algorithm. We'll look at how features are extracted and updated for algorithms like Support Vector Machines The human emotion dataset can be used to investigate the robustness and design of classification algorithms, as well as how they behave for various dataset types.

Face recognition algorithms are usually applied to the image or captured frame before extracting features for emotion detection. The (FACS) Facial Action Coding System assigns a numerical value to each facial moment. - of these numbers is referred to as an action unit. The action unit-based facial action coding system is a good way to find out which facial muscles are involved in which speech. Face landmarks are extremely significant and can be used to detect and recognize faces. Good features are those that aid in correctly identifying the object. These features and expressions can be used to build real-time face models. A local feature is made up of feature descriptors. A histogram is an example of a function descriptor. [8]

3.6 Company Verification

The skill set of a specific candidate is fetched from his or her resume-by-resume parsing. On Naukari.com, there are several profiles; DNN is applied to these profiles, and a model is developed and applied.

We can quickly determine which job is best for an applicant with a specific skill set using the model we've generated.

Kera's is an effective library which helps in developing deep learning models and it is simple to use. [8] The DNN algorithm is applied to a total of 25 classes (job recommendation roles) and 10,000 rows, assisting the organization in deciding if an applicant is appropriate for a specific work.

DATA PROCESSING

4.1 Data collection

The interviewees' answers, including both sound and visual data, were recorded for examination. The applicants were instructed that the total with respect to their screenings and responses, including sound and video, would be recorded and separated according to our observations and used as references" for recruiting suggestions.

The inquiries during the AVI were organized in a standard way. Every one of the candidates were given similar [9] five inquiries, which were typically arranged to survey the candidates' relational abilities dependent on the expected set of responsibilities.

4.2 Data Labelling

The distance between facial features is measured with this mapping. This recognizes facial landmarks that identify the face and allows for the development of a facial signature. [9] A camera searches for the distinguishing features of a face and algorithms assist in determining the orientation and movement of the face.
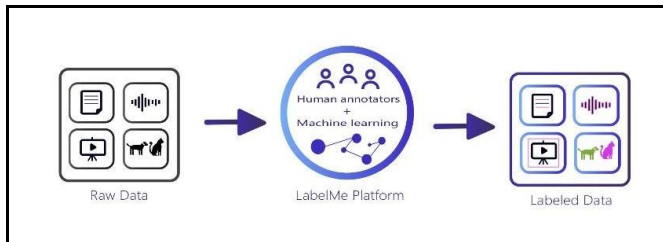

Fig 4: Data Labelling → Labelled Emotions

We will also take into consideration the application in Film Industry Sector and also how we can modify our recruitment system into Asynchronous [10] AI Auditioning System for a particular role in a film.

4.3 Feature Extraction

To catch the candidates' outward appearances, we began with the pretrained Inception-v3 dataset gathered for ImageNet, which incorporates in excess of 14 million pictures assembled into 1,000 classes.
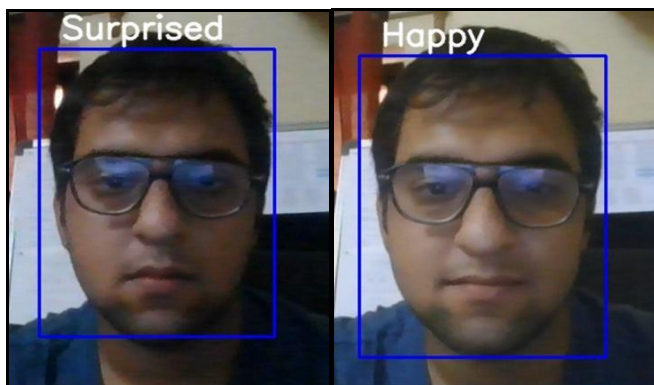

Fig 5: Facial Emotion

The width of the relative multitude of pictures was standardized to 320 pixels, while the structure of picture was dictated by the pixel of the picture gadget.

To improve picture portrayal and reduce establishment impedance from hair and features, we changed all of the photos over to grayscale. The experiments utilized in this investigation contained in excess of 10,000 pictures.

MODEL BUILDING

DNN Model was used for resume parsing and verification. In this, we collected 25 Job roles data from Naukri.com which contains ten thousand rows of Data. It parses the resume in pdf file to sort and recommend the best job role using machine learning. [11]
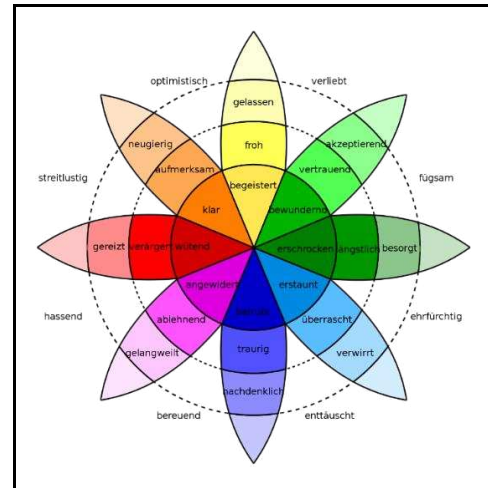

Fig 6: Theory of Emotions by Psychologist Robert Plutchik

For Voice Emotion Recognition, we have used approximately 1GB of raw data from Kaggle which consists of 2453 audio files. For Face Emotion recognition training we have used dataset from Kaggle which is classified into 7 face emotions which are angry, disgust, fear, happy, neutral, sad and surprised. There are a total of 21,000 images in this dataset.

The development plan consists of many steps for the development diagram. Firstly, the dataset of audio files is uploaded on google drive. Then Jupiter notebook in google Collab is used to load the dataset from the drive and train the CNN with TensorFlow and keras to get the predicted emotions

In Toward the simulation of emotion in synthetic speech, Murray and Arnott attempted to recognize emotions on the basis of rate, pitch average, pitch range, intensity, voice quality, voice changes and pitch of the speech
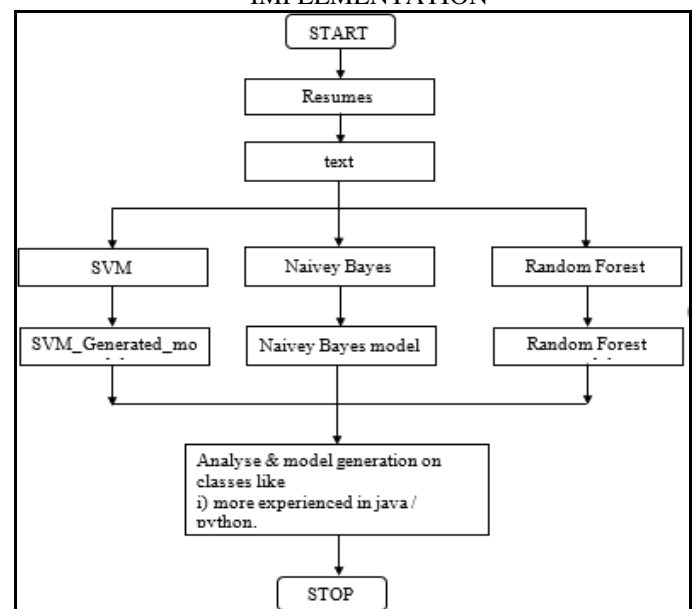
IMPLEMENTATION

Fig 7: Flowchart Architecture

Implementation consolidates all of those activities that happen to change over from the old system to the new. The old system includes manual exercises, which is worked in an entirely unexpected path from the proposed new structure. A suitable execution is basic to give a reliable structure to meet the necessities of the relationship There are a couple of methodologies for managing the execution and the resulting change from the old to the new electronic system. [12]

The most secure procedure for change from the old system to the new structure is to run the old and new system in equivalent. In this strategy, an individual may work in the manual more prepared planning structure similarly as start working the new automated system. This technique offers high security, in light of the fact that whether or not there is an imperfection in the electronic system, we can depend on the manual structure. In any case, the cost for keeping two systems in equivalent is incredibly high. This surpasses its benefits.

Another overall procedure is a prompt cut over from the current manual structure to the modernized system. There are no equivalent activities. This framework requires wary arranging. [13] All the task will be done by our gathering and working structure will be made by us each individual will be designated with the work. also, when required. Be that as it may, this procedure is less ideal due to the inadequacy of total of the system.

RESULTS

The projected successfully parsed the resume with job recommendation taken from 25 job roles, by parsing the resume and with the use of machine learning CNN model a successful job recommendation is provided.
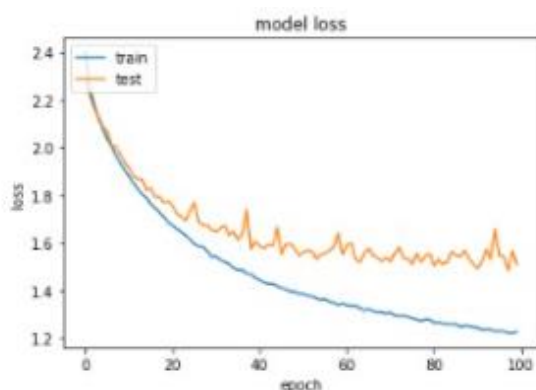

Fig 8: Speech Recognition

In the speech recognition the model is successful to Identify the voice of male or female and also classify it into different emotions such as sad or angry by the tone of the speaker with its initial accuracy of 0.5500 to accuracy of 0.9150 after training of the initial data the initial loss of data is 1.2252 and after training and testing of the data the loss has come down to 0.1619. [14]
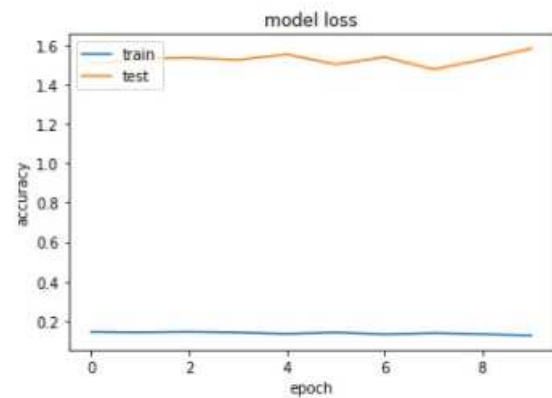

Fig 9: Facial Emotion Recognition

In the video emotion recognition, the system is successful to identify different emotions such as happy, sad, disgusted or angry which is useful for automatic job recommendation.

The initial accuracy of the system was 0.2560 which has increased to 0.9476 the initial loss of data was 1.8090 which has been reduced after testing and training of the data to 0.1471. The model is successful to provide job recommendation and assist the company in the recruitment project.

FUTURE DIRECTIONS

This paper built up a project with a TensorFlow-based Deep Learning model to precisely perceive a candidate's actual character dependent on just 100 genuine examples of occupation candidates. Our approach accomplished an exactness above 93%, beating past related research center examinations whose precision ran somewhere in the range of 64% and 75% with regards to nonverbal correspondence.

The APR used in this AVI can be embraced to enhance our character evaluation techniques that can be contorted by work candidates because of the impacts of social craving to be chosen for organizations. [15]

Past related work has discovered that highlights learned by profound neural organizations can convey better exhibitions in foreseeing the enormous five characteristics than can highlights. In future work, we may join our methodology with highlights to figure out how to perceive a candidate's character. Future exploration ought to incorporate a more different member populace. [15]

CONCLUSION

This project is for character computing. In customary character figuring, approving APR utilizing physically marked highlights from any conceivable recognizable distal prompts was very convoluted. This undertaking built installed with a Tensor Flow-based supervised DL model to precisely interviewee's actual occupation candidates. Our APR approach accomplished an exactness above 90%, beating past related lab examines whose precision ran somewhere in the range of 61% and 75% with regards to nonverbal correspondence. [16]

The high-performing APR utilized in this paper can be embraced to enhance or supplant self-announced character appraisal techniques that can be twisted by work candidates because of the impacts of social craving to be chosen for business.

Past related investigations have discovered that multimodal highlights (picture casings and sound) learned by profound neural organizations can convey better exhibitions in foreseeing the huge five characteristics than can unimodal highlights.

ACKNOWLEDGMENT

REFERENCES

[1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph et al., "A view of cloud computing," Communications of the ACM. 1999

[2] R. A. Popa, F. H. Li, and N. Zeldovich, "An ideal-security protocol for order-preserving encoding," in Proceedings of the 2013 IEEE Symposium on Security and Privacy (SP'13). IEEE, pp. 463–477, 2013.

[3] KaipGing Xue, ShaoFhua Li, JianDan Hong, Yingjie Xue, Nenghai Yu, and Peilnin Hong "Two-Cloud Secure Database for Numeric-Related SQL Range Queries with Privacy Preserving", 2020

[4] J.W. Rittinghouse and J. F. Ran some, Cloud computing: implementation, management, and security. CRC press, 2018.

[5] D. Boneh, D. Gupta, I. Mironov, and A. Sahai, "Hosting services on an untrusted cloud," in Advances in Cryptology EUROCRYPT 2015. Springer, pp. 404–436, 2015.

[6] J.-M. Bofhli, N. Gruschka, M. Jensen, L. L. Iacono, and N. Marnau,"Security and privacy-enhancing multicloud architectures," IEEE Transactions on Dependable and Secure Computing, vol. 10, no. 4, pp. 212–224, 2013.

[7] K. Xufe and P. Honng, "A dynamics secure group sharing framework in public cloud computing," 2018

[8] D. Zissis and D. Lekkas, "Addressing cloud and machine learning " CRC press , 2020

[9] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph et al., "A view of cloud computing", Communications of the ACM, vol. 53, no. 4, pp. 50–58,2010.

[10] Y. Yang, H. Li, M. Wen, H. Luo, and R. Lu, "Achieving ranked range query in smart grid auction market," in 2014 IEEE International Conference on Communications (ICC2014). IEEE, Vol.2, No.4,April 2014

[11] C. Wang, Q. Waxng, K. Rcen, N. Ccao, and W. Lou, "Dependable storage services in cloud computing", IEEE Transactions on Services Computing, vol. 8, no. 2, pp. 220–202, 2014

[12] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu, "Order preserving encryption for numeric data," in Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data. ACM, pp.563–574, 2004.

[13] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, "Toward secure and dependable storage services in cloud computing," 2008

[14] D. Zissfsis and D. Lekkcas,"Addressing cloud computing security issues", Future Generation Computer Systems, vol. 8, no. 8, pp. 503–562, 2019.

[15] R. A. Popnza, C. Redfvield, N. Zeldofvich, and H. Balakrishnan, "Crypt: protecting confidentiality with encrypted query processing," in Proceedings of the 20th ACM Operating Systems Principles. ACM, pp. 85–190, 2011.

[16] F. Hao, J. Daugman, and P. Zielinski, "A fast search algorithm for a large fuzzy database," IEEE Transactions on Information Forensics and Security, vol. 3, no. 2, pp. 203– 212, 2008.