Blockchain Based Career Recommendation System through Data Mining utilizing Cryptocurrency

Jomer Catipon [1], Eliniel Valdez[2], Ganmarie Duque[3], Jianna Elisha Florendo[4], Glenn Marvin Del Rosario[5], Charis Dane Esplana[6], August Thio-Ac[7], Mark Melegrito[8], Glenn Virrey[9], Jessica Velasco[10], Lean Tolentino[11]

1,2,3,4,5,6,7,8,9,10,11 *Electronics Engineering Department, Technological University of the Philippines, Philippines.*

Abstract— **This project successfully developed and deployed Career Oracle, a web-based recommendation system that provides ECE Graduates and Graduating students' specialization recommendations with blockchain encryption that uses Cryptocurrency as proof-of-concept payment to undergo and verify the transactions/notarization/encryption. The web-application contains the database of grades, technical-assessment test results, behavioral-assessment test results as well as Board Exam Ratings that will be fitted on the machine learning model to provide recommendations. Recommendations were provided through the use of Ensemble Machine Learning Model which used the data of ECE Graduates that are currently working. To verify the validity and reliability of the model, 15 Graduating Students and 15 Graduate Students were asked to answer an evaluation based on the Likert's Scale with a result of 4.41, falling under the "Extremely Effective" on the scale. Aside from the validity of the model, website appeal and website security were also evaluated with a result of 4.49 and 4.46 respectively. Aside from the user-based validation, the recommendation algorithm was evaluated through root mean square error (RMSE) and mean square error (MSE), the result was 0.57 and 0.326 respectively. However, the system falls behind in RMSE which is 14% above the accepted error threshold of 0.2-0.5. Lastly, the notarization of the certificates generated after the recommendation was encrypted and notarized through NEM Apostille Function. For the notarization hash to go through the NEM Blockchain, the development of Cryptocurrency (TUPCoin) was done. The said Cryptocurrency token was successfully developed and deployed on the Binance Test Net Smart Chain.**

Keywords— **Ensemble Machine Learning, Cryptocurrency, NEM Blockchain, Token, Binance, Smart Chain, Recommendation System, RMSE, MSE.**

I. INTRODUCTION

ECE encompasses a variety of careers/specializations. In fact, ECE graduates and graduating students, deal with job-mismatches because of the lackluster in information regarding their current strengths, wants and preferences. Blockchain technology provides transparency and proof that a transaction happened, because of this, providing credibility on certificates and information serves to be an easy task.

Aside from being primarily used to be a store of value and a ledger system for financial transaction, Blockchain technology can also be used to verify the content of a certificate through its hash key and transaction hash, given that it is processed through the blockchain. Furthermore, payments can also be done through using cryptocurrencies, making the process and transaction be verified on two blockchains, providing a more secure and reliable certificate.

Recommendation Systems have evolved from using Naked Predictive Analytics to Machine Learning Models. This is because machine learning automates predictive modelling by automatically looking for patterns and behaviors in the data [1]. However, most recommendation system fails at the start because of cold-start problems [2]. To provide solutions, ensemble machine learning models through gradient boosting, bagging and stacking can be used. Aside from ensemble models, Domain adaptation can also be used.[3] [4]

The recommendation system used in the web application, leverages gradient boosted ensemble models to provide recommendations. The model endpoint was deployed on AWS Sagemaker and trained on the same platform.

The research project mainly focuses on the integration of the ensemble machine learning model, blockchain technology, blockchain notarization and cryptocurrencies as proof-of-concept-payment. Hence, this will not tackle any specific details regarding the data analysis, visualization and actual training data and test-data for the recommendation system.

The paper is presented as follows: Section II defines the methods done by the proponents to leverage and access NEM Apostille Function. Furthermore, Section II also defines the methods used to develop the model and cryptocurrencies. Section III involves the evaluation results of the website, security and reliability of the model. Section IV states the conclusion while Section V provides recommendation for possible future works of the research.
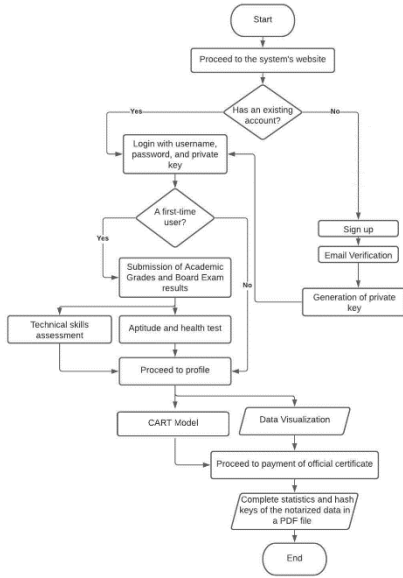
II. METHODOLOGY

Figure 1. System Architecture

Figure 1 shows the flow and system architecture for navigating the web application
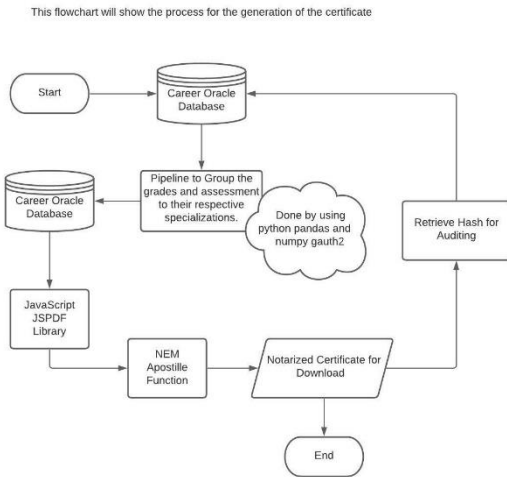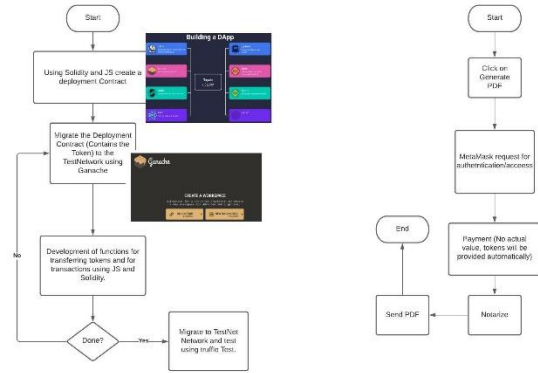

Figure 2. PDF, NEM notarization

Figure 3. Cryptocurrency Development

Figure 2 shows the flow for the notarization of the certificate through NEM and the generation of the certificate through JS libraries CanvasJS and JSPDF.

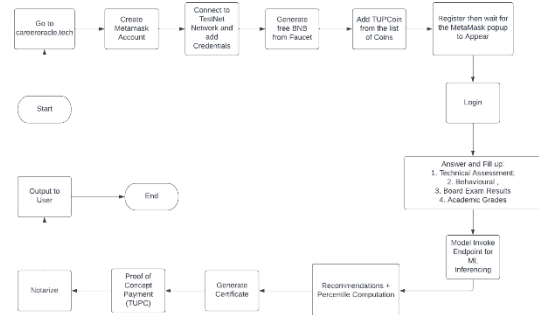Figure 3 illustrates the process for the development of the cryptocurrency (TUPC) as well as how it was employed.

Figure 4. Software Cycle and Whole process

Figure 4 illustrates the actual flow for the recommendation to work. The Ensemble Machine Learning Model was trained on AWS Sagemaker and deployed on the same Web Service with Machine Learning Model Endpoints for Lambda Inferencing. The API Gateway was also developed on the same platform. The front-end system will invoke the API url and transfer the "uid" of the graduate that is stored on the cloud-firestore database. This UID will be the key to provide data to be processed through a micro ETL pipeline which will then be used to be fed on the ML Endpoint for actual recommendation.

III. RESULTS AND DISCUSSION

A PDF file consists of different specialization recommendation from the web application that also contains the HASH key to be verified on the NEM Blockchain. A Sample certificate is provided on Figure 5.
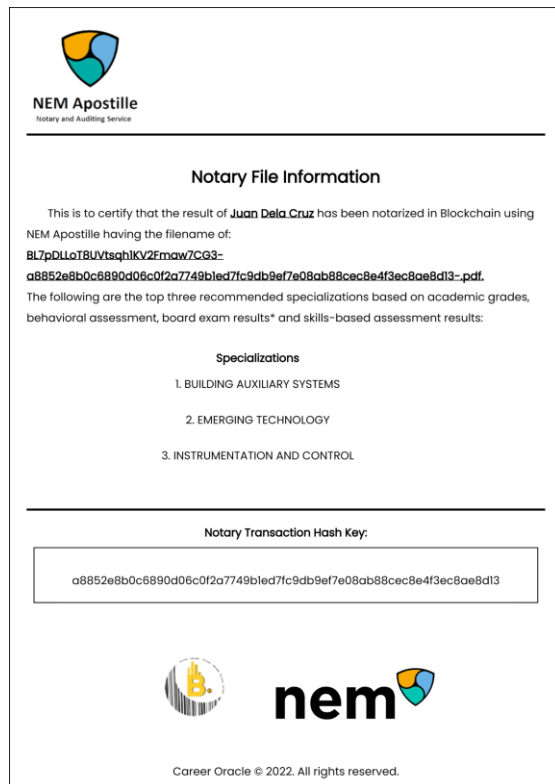
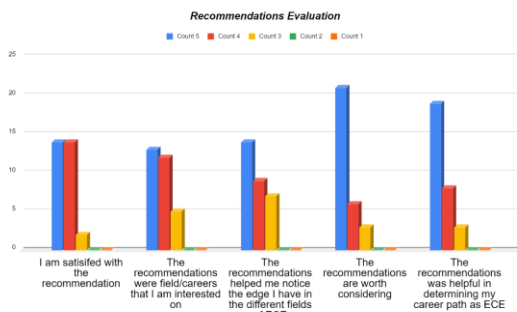**Figure 5.** Sample Certificate with notarization and recommendations
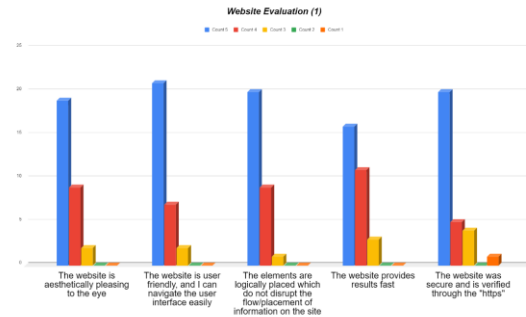

Figure 6. Recommendation Evaluation

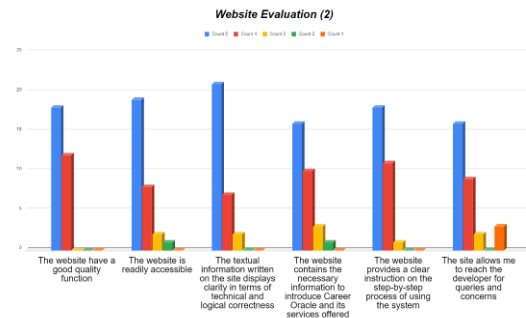

Figure 7. Website Evaluation (1)

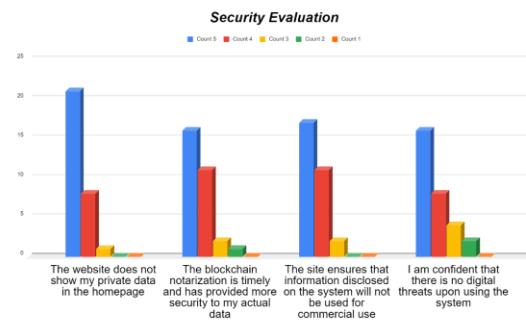**Figure 8.** Website Evaluation (2)
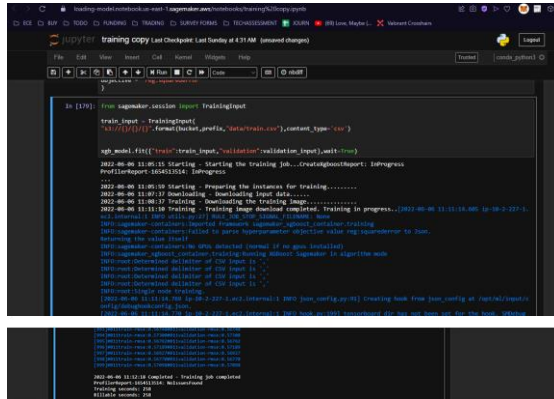

Figure 9. Security Evaluation

Figure 10. Model Training and Evaluation through AWS Sagemaker

Figure 6 to figure 9 shows the summary of evaluation on the recommendation, website and security respectively. 30 respondents which consists of 15 ECE Graduates and 15 ECE Graduating students tried the system, evaluated and gave their respective responses in accordance with ISO 9126 Likert's Acceptability Scale. The average rating through all the evaluation falls under the "Extremely Satisfied" scale on the Likert's Scale.

Figure 10 shows the proof of training through a code snippet on AWS Sagemaker as well as the model evaluation.

IV. CONCLUSIONS

The generated recommendation result from the system were successfully notarized and registered through the NEM Blockchain which provides credibility for the certificate as well as the data provided. The model was evaluated with an objective of RMSE and eventually had an error of 0.57 which is 14% slightly above the threshold 0.2 - 0.5. Furthermore, the development of the TUPC for proof-of-concept payment was successfully accomplished and deployed on Binance Smart Chain.
The respondents were "Extremely Satisfied" with the system with respect to the recommendation results, the website UI as well as the security. Furthermore, the development of an ensemble machine learning model was successfully done and deployed on the cloud using AWS Sagemaker.

V. FUTURE WORKS

The project may be developed further to address more the problem regarding cold start by implementing the recently published journal about "Infinite Recommendation Networks: A Data-Centric approach". Furthermore, the system could also have a learning section where the ECE Graduates and Graduating students can retrieve/procure the right documents/learning materials for effective learning to happen and jumpstart their career.TUPC for proof-of-concept payment was successfully accomplished and deployed on Binance Smart Chain.
The respondents were "Extremely Satisfied" with the system with respect to the recommendation results, the website UI as well as the security. Furthermore, the development of an ensemble machine learning model was successfully done and deployed on the cloud using AWS Sagemaker.

REFERENCES

[1] K. Wakefield. *Predictive Modeling Analytics and Machine Learning*. [Online]. Available: https://www.sas.com/en_gb/insights/articles/analytics/a-guide-to-predictive-analytics-and-machine-learning.html

[2] V. Lendave. (2021). *Cold-Start Problem in Recommender Systems and its Mitigation Techniques*. [Online]. Available: https://analyticsindiamag.com/cold-start-problem-in-recommender-systems-and-its-mitigation-techniques/

[3] (2021). Open Data Science website. [Online]. Available: https://odsc.medium.com/overcoming-the-cold-start-problem-how-to-make-new-tasks-tractable-b43ab79bf512

[4] S. Najafi and Z. Salam, "Evaluating Prediction Accuracy for Collaborative Filtering Algorithms in Recommender Systems", Computer Science thesis, KTH Royal Institute of Technology, Stockholm, Sweden, May 2016.

[5] A. Asgarian. Open Data Science Conference website. [Online]. Available: https://odsc.com/speakers/overcoming-the-cold-start-problem-how-to-make-new-tasks-tractable/?__hstc=39712252.1659f1977841582758c48a740b259d0a.1592859022142.1619044727420.1619096932918.123&__hssc=39712252.45.1619096932918&__hsfp=2561499225