Machine Learning for Solar Panel Fault Detection

Sanika Naik, Glen Uehara, Sameeskha Katoch, Dr. Andreas Spanias

***Abstract*—With the world's growing energy crisis becoming a more prevalent issue, solar energy has risen as the leading sustainable and cost-effective replacement for fossil fuels. As is the case with all emerging industries, progress is accompanied by new barriers that must be addressed. In this case, the efficiency of solar panels requires constant monitoring of the voltage, current, temperature and irradiance. This project aims to recognize faults and classify them as soiling, degraded modules, shading or arc faults by characterizing these features as the inputs in artificial neural network models then comparing the results with those of a quantum model.**

Keywords— machine learning, fault detection, logistic regression model, support vector machines, neural networks, classical model, quantum model

I. Introduction

The constantly increasing nature of the world's demand for energy coupled with limited natural resources results in a growing need for renewable energy sources. Approximately 1.8×10^{14} kW of solar energy is intercepted by the Earth, making it the most abundant source of renewable energy [1]. Efficient photovoltaic systems have the potential to serve as a viable solution to the energy crisis.

PV technology directly converts sunlight into electricity and the power generated by the individual cells is determined by the intensity of the light. Despite the statistic that PV installations currently provide only 0.1% of the world's total electricity generation, PV technology is predicted to deliver around 1081 GW by 2030 [2]. This significant measurement of growth further establishes the need for development in monitoring and fault detection technologies to minimize the solar energy wasted in the process. Failure to do so has the potential to compromise the entire solar array system.

While multiple faults can occur, this project will specifically focus on detecting soiling. This fault results from particles like dirt, snow or dust that cover the surface of PV modules and obstruct the amount of solar energy delivered. The solution to this problem is the consistent cleaning of the modules. Studies have shown that annual losses of soiling range from 1.5% to 6.2% depending on the location of the plant [3].

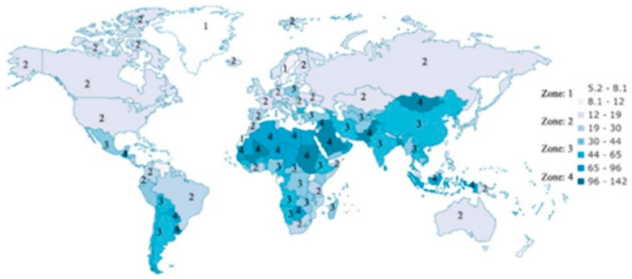

Figure 1: dust intensity around the world [3]

One of the main difficulties of detecting this fault stems from the immense amount of data that must be analyzed. Faults are identified by a combination of various measurements and must be properly diagnosed. For example, soiling will result in measured irradiance aligned with standard testing conditions, however measured power produced will be significantly lower [4]. Machine learning systems allow us to recognize patterns in this data and are 24.6 percent more precise than human inspected methods [5]. In this project, I will run the data through a logistic regression, support vector machine and neural network models to analyze test data and determine which is most accurate for the data set provided. Then, I will evaluate those results in comparison with the quantum version of that specific model. Quantum algorithms have the potential to drastically increase our powers of computing and solve problems of a much greater scale. However, currently this field is undergoing much progress and quantum algorithms are not yet ideal for all types of data. The purpose of this comparison is to determine whether quantum algorithms are suitable for addressing this problem.

II. Data

The simulated data used in this research went through multiple cleaning and formatting processes prior to being inputted into the regression model. We started by loading the solar data into Google Colab in arrays X and Y. Next, we separated the data from the labels and converted it from a Pandas data frame to a NumPy array. We then converted the one-hot encoding labels into categorical labels for ease of use and normalized the data to serve as a viable input into the model. Finally, the data was divided into a 70/15/15 train/validation/test split.

III. Logistic Regression

This project first trained SKLearn's Logistic Regression model to detect solar soiling faults in the system. This specific model performs at a relatively faster speed than other supervised classification techniques, however, it can tend to be less accurate due to its simplistic nature [6]. This model was optimized using the 'saga' solver with 5000 maximum iterations to produce a test accuracy of 97.94%.

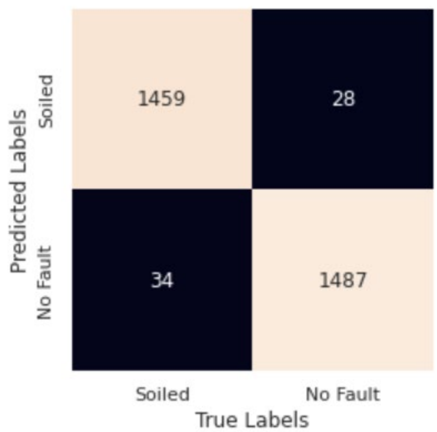

Figure 2: logistic regression confusion matrix

IV. Support Vector Machines

This project then trained SKLearn's Support Vector Machine model to detect solar soiling faults in the system. This model separates two classes by creating a decision boundary and aims to maximize the margin between this hyperplane and data [7]. While this method has its advantages, it is not suitable for large databases due to their greater amounts of noise [8]. To optimize this model, we held break ties as true and used the ‘linear’ kernel with 5000 maximum iterations. This yielded a test accuracy of 98.17% as displayed in the confusion matrix below.

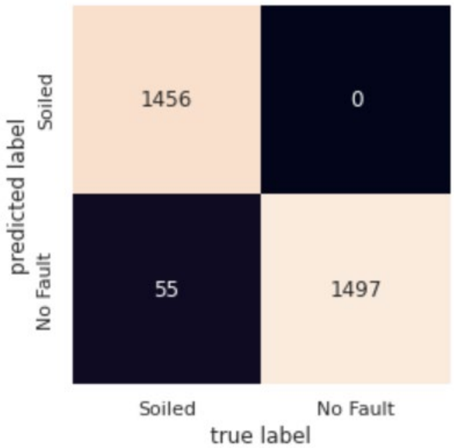

Figure 3: support vector machine confusion matrix

V. Artificial Neural Network

The classical portion of this project then trained SKLearn’s Artificial Neural Network model to detect solar soiling faults in the system. This model required a higher level of training due to its susceptibility to overfitting [9]. We used a one layer neural network with 250 hidden nodes as our model structure then optimized it using the logistic activation function, the ‘adam’ solver and 100 maximum iterations. Finally, the model resulted in the highest test accuracy of 98.34%.

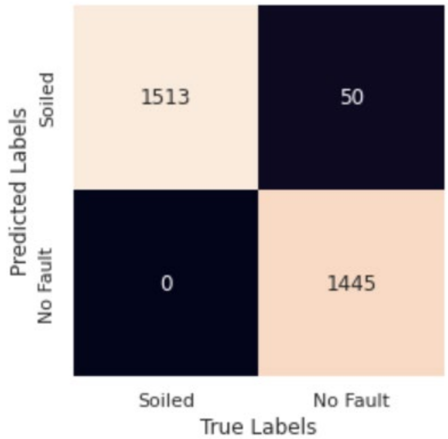

Figure 4: artificial neural network confusion matrix

VI. Quantum Neural Network Model

In theory, quantum computers are capable of solving certain types of problems at a faster speed than classical computers. Researchers are working to determine whether this is the case in the field of machine learning where a computer is trained to solve practically relevant problems. Preliminary studies have begun to show that, when compared to classical neural networks, quantum models were able to achieve higher effective dimensions [10]. This can be observed in Figure 5 below.

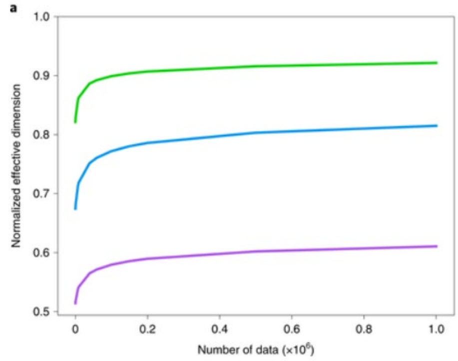

Figure 5: normalized effective dimension [10]

Effective dimensions are an essential indicator of how trainable a neural network model is and measures how capable a model is of adapting to new data [10]. Thus, a higher proportion of optimized parameters were being actively used in the quantum neural network and lowering the training loss as compared to other models.

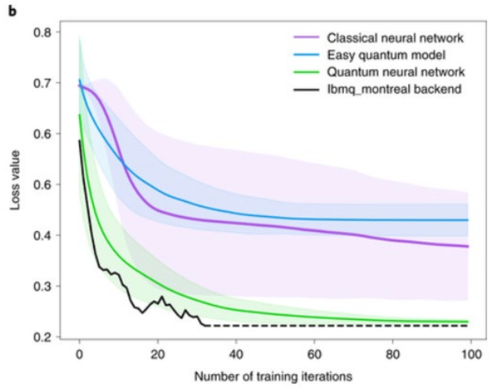

Figure 6: training loss [10]

In this project, we wanted to expand upon these results and conduct our own research to determine whether there was a quantum advantage to neural network models.

We used a four layer neural network with 50 epochs and a learning rate of 0.01. The first was a conventional neural network layer with 10 inputs and 4 outputs. This was followed by a two layer quantum neural network and the last layer was a conventional layer with 1 input and 2 outputs. Lastly, a sigmoid activation function was applied to the data. This optimization resulted in a 61.3% accuracy as displayed in the confusion matrix below.

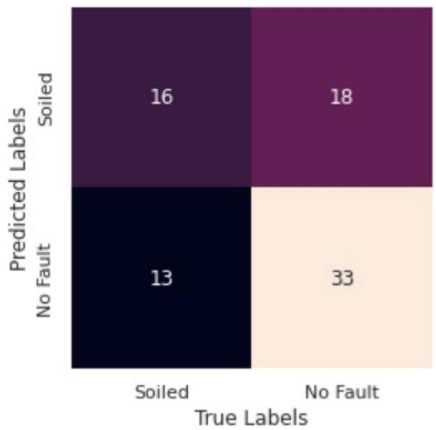

Figure 7: quantum neural network confusion matrix

It should be noted, however, that past research in this field produced successful quantum results of ~90% accuracy which is significantly higher than the results produced in this project [11]. Thus, we know that our main issue came from the optimization of the algorithm rather than the failure of quantum neural networks to address this issue.

VII. Quantum Support Vector Machines Model

The relatively lower results we faced with the quantum neural network model lead us to create a quantum SVM model for comparison. Quantum SVM models work by converting the classical data into quantum states using quantum feature maps and then build a kernel using the quantum states [12]. This kernel matrix allows the quantum model to be trained in the same way as the classical model.

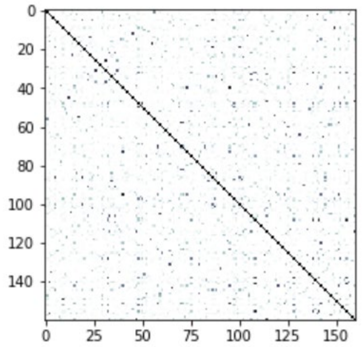

Figure 8: SVM kernel

This model yielded more accurate results- specifically a test accuracy of 97.5% as displayed in the confusion matrix below.

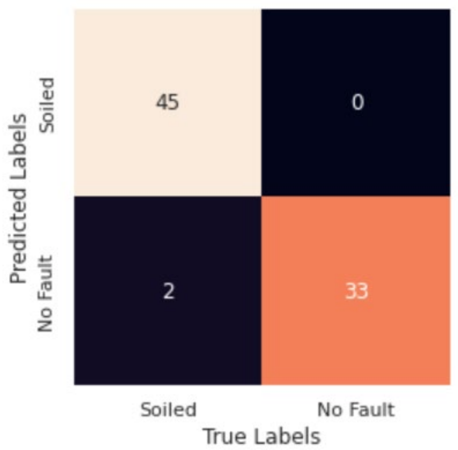

Figure 9: quantum neural network confusion matrix

VIII. Classical Versus Quantum Results

When comparing classical and quantum results, a significant difference can be observed. There were a number of challenges with optimizing the quantum machine learning algorithms and aspects that made them inefficient as the primary method of addressing this problem. QML algorithms required more processing power and had longer execution times. Furthermore, we were unable to run the full dataset without the program crashing.

Prior research in the field has shown that quantum SVM showed a significant advantage over classical SVM for multi-class classification problems but not for binary classification problems [13]. This research project reached the same conclusion and showed similar results for classical and quantum SVM algorithms.

In terms of quantum neural networks, prior research allows us to conclude that quantum neural networks have the potential to yield higher results than those of this study. Thus, quantum neural networks have the capability of producing a higher accuracy than classical models but there are still limitations on our ability to mimic these results.

IX. Conclusion and Furture Work

Preliminary quantum neural networks results indicated a need for improvement and further optimization. Potential future work would be to continue to improve the parameters of the model and aim for a higher accuracy. Furthermore, this project was predominantly run on Google Colab and would have been benefitted by being run on an actual quantum computer. This is something we hope to do in the future in order to lessen the run time and determine the potentially higher accuracy this will yield.

Acknowledgment

This project was sponsored by National Science Foundation Award 1854273.

References

[1] N. Kannan and D. Vakeesan, “Solar Energy for future world: - A Review,” *Renewable and Sustainable Energy Reviews*, vol. 62, pp. 1092–1105, Sep. 2016.

[2] V. V. Tyagi, N. A. A. Rahim, N. A. Rahim, and J. A. L. Selvaraj, “Progress in solar PV technology: Research and achievement,” *Renewable and Sustainable Energy Reviews*, vol. 20, pp. 443–461, Sep. 2013.

[3] M. R. Maghami, H. Hizam, C. Gomes, M. A. Radzi, M. I. Rezadad, and S. Hajighorbani, “Power loss due to soiling on Solar Panel: A Review,” *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 1307–1316, Jan. 2016.

[4] S. Rao, S. Katoch, and V. Narayanaswamy, *Machine learning for solar array monitoring, optimization, and Control*. San Rafael, CA: Morgan & Claypool Publishers, 2020.

[5] J. A. Dhanraj, A. Mostafaeipour, K. Velmurugan, K. Techato, P. K. Chaurasiya, J. M. Solomon, A. Gopalan, and K. Phoungthong, “An effective evaluation on fault detection in solar panels,” *Energies*, vol. 14, no. 22, p. 7770, 2021

[6] A. C. Chang, “Machine and deep learning,” *Intelligence-Based Medicine*, pp. 67–140, 2020.

[7] “Applications of support vector machine (SVM) learning in cancer genomics,” *Cancer Genomics & Proteomics*, vol. 15, no. 1, 2018.

[8] R. Gholami and N. Fakhari, “Support Vector Machine: Principles, parameters, and applications,” *Handbook of Neural Computation*, pp. 515–535, 2017.

[9] S. Dreiseitl and L. Ohno-Machado, “Logistic regression and Artificial Neural Network Classification models: A methodology review,” *Journal of Biomedical Informatics*, vol. 35, no. 5-6, pp. 352–359, 2002.

[10] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, “The power of Quantum Neural Networks,” *Nature Computational Science*, vol. 1, no. 6, pp. 403–409, Jun. 2021.

[11] G. Uehara, A. Spanias, W. Clark, “Quantum Information Processing Algorithms with Emphasis on Machine Learning,” *Proc. IEEE IISA 2021*, July 2021.

[12] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta,

"Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, no. 7747, pp. 209–212, Jun. 2018.

[13] Havenstein, Christopher; Thomas, Damarcus; and Chandrasekaran, Swami (2018) "Comparisons of Performance between Quantum and Classical Machine Learning," *SMU Data Science Review*: Vol. 1: No. 4, Article 11.

[14] Glen Uehara, Sunil Rao, Mathew Dobson, Cihan Tepedelenlioglu and Andreas Spanias, "Quantum Neural Network Parameter Estimation for Photovoltaic Fault," *Proc. IEEE IISA 2021*, July 2021

[15] Kristen Jaskie and Andreas Spanias, Positive Unlabeled Learning, Morgan & Claypool Publishers, AI and Machine Learning Eds. R. Brachman and (Cornell) and F, Rossi (IBM Research AI), 152 pages, ISBN 9781636393087, February 2022.

[16] A. Dixit and U. Shanthamallu, A. Spanias, V. Berisha, M. Banavar, Online Machine Learning Experiments in HTML5, *IEEE FIE 2018,* San Jose, October 2018

[17] G. Uehara, V. Narayanaswamy, C. Tepedelenlioglu, A. Spanias, *"Machine Learning for Photovoltaic Topology Optimization,"* 2022 IEEE 13th International Conference on Information, Intelligence, Systems & Applications (IISA), Hybrid Conference, July 2022.

[18] M. Yarter, G. Uehara, A. Spanias, *"Implementation and Analysis of Quantum Homomorphic Encryption,"* 2022 IEEE 13th International Conference on Information, Intelligence, Systems & Applications (IISA), Hybrid Conference, July 2022.

[19] H. Braun, S. Buddha, V. Krishnan, C. Tepedelenlioglu, A. Spanias, S.i Takada, T. Takehara, M. Banavar, and T. Yeider., Signal Processing for Solar Array Monitoring, Fault Detection, and Optimization, Synthesis Lectures on Power Electronics, Morgan & Claypool, Book, 1-111 pages, ISBN 978-1608459483, Sep. 2012.