Presented by: Andrew Gray, Shahar Journo, Scott Jue, Johanna Sheu
MSDS 498 Capstone Project, Sec 57
Spring 2023

SOLAR TECH SOLUTIONS

EXECUTIVE SUMMARY

TABLE OF CONTENTS

INTRODUCTION

As the associated costs with solar cells decrease, the demand for solar energy is rapidly increasing. To meet these demands, energy utilities need to find ways to maximize their use of renewable energy in the grid. Not only does renewable energy combat climate change, but it also represents lower costs for the utility companies and lower energy prices for consumers. However, solar energy poses a unique challenge for utilities: once captured it cannot be reliably stored and must be fed into the grid immediately. Due to this dynamic, it is important that energy utilities understand the potential solar capacity of an area and how much of that capacity is being captured through residential solar panels to make strategic decisions about grid utilization.

MISSION AND GOALS

Solar Tech Solutions provides a predictive tool using weather data by location for electric utility companies to forecast Photovoltaic (PV) energy production with competitive accuracy. This allows for the right balance and allocation of renewable energy resources and available fossil fuel resources to power the grid. Municipalities can also plan to build infrastructure according to the predicted rate to adequately capture the increasing supply of solar energy.

BUSINESS OPPORTUNITY

This project will allow for the optimization of solar panels as part of a municipality’s electric power grid to reduce unnecessary usage of fossil fuels, and if implemented correctly will represent a cost savings for both the customers and the utility company.

BACKGROUND

Electric utility **customer satisfaction** depends on continuous delivery (no blackouts), low cost, and environmental impact of energy production. Electric utility **profitability** depends on minimizing supply and associated cost of energy fed into the grid while continuously meeting demand for energy from their customers.

Utility companies have a mix of production methods each with its own cost and environmental impact. If fossil fuel energy production can be shut off when solar production is most high, that represents increased profitability for the utility company, and increased customer satisfaction.

Weather and solar panel capacity are the primary factors that determine solar power’s contribution to the grid, so it would be beneficial for a utility company to have a highly accurate model so they can make quick decisions regarding what production mix they need to use to meet demand.

OBJECTIVES

Improve forecast accuracy: machine learning algorithms can improve the accuracy of solar energy output predictions compared to traditional methods. This can help utility companies better plan for and manage energy demand and supply, leading to more efficient and cost-effective operations.

Cost savings: predicting solar energy output can reduce the need for expensive backup power sources and avoid overproduction of energy, resulting in cost savings for utility companies.

Scalability: machine learning algorithms can be trained on large amounts of historical and real-time data, making them scalable to different types of solar energy systems and geographies.

Competitive advantage: by adopting a machine learning approach, utility companies can gain a competitive advantage in the energy market and improve their reputation as a forward-thinking and innovative organization.

Environmental benefits: solar energy is a clean and renewable energy source, and by using machine learning to optimize the production and use of solar energy, utilities can reduce their carbon footprint and contribute to a more sustainable future.

AUDIENCE AND MARKET

Electric utility companies located in the northern hemisphere that utilize solar panels as part of their mix of energy sources.

Government agencies and municipalities interested in integrating solar power into existing power grids.

Solar panel companies that wish to develop a feasibility study to expand operations.

Utility customers who may want to install solar panels based on long-term cost savings.

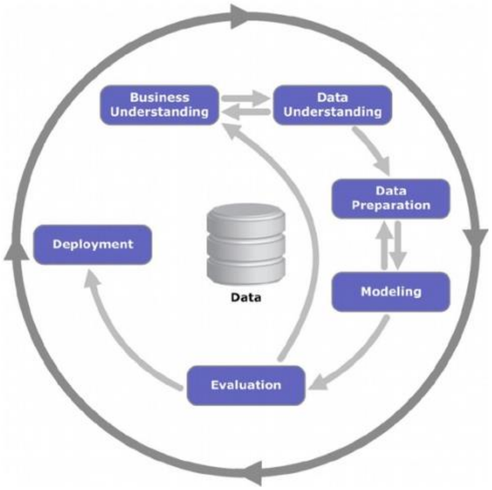
METHODOLOGY

PROCESS MANAGEMENT

We will use the CRISP-DM approach to develop and implement an analytical solution for electric utility companies.

The process involves six sequential phases which include:

1. Business understanding – determine business objectives, assess situation, determine data mining goals, produce project plan.
2. Data understanding – acquire necessary data, describe data, explore data, verify data quality
3. Data preparation – select, clean, construct, integrate and format data.
4. Modeling – select modeling techniques, generate test design, build model, assess model.
5. Evaluation – evaluate results, review process, determine next steps.
6. Deployment – plan deployment, plan monitoring and maintenance, produce final report, review project.



DATASET

The dataset publicly sourced from Kaggle contains solar power output data from solar panels located at 12 Northern hemisphere sites over a 14-month period between 2017 and 2018.

The independent variables include:

- **Numerical**: altitude, humidity, ambient temperature, power output from the solar panel, wind speed, visibility, pressure, cloud ceiling
- **Categorical**: season
- **Text**: date, time sampled, location

We also used data from Solcast, an online resource of solar assessment and forecasting data, and merged it with the Kaggle dataset for additional weather features.

The variables from the Solcast dataset used include:

- **Numerical**: CloudOpacity, DewpointTemp, PrecipitableWater, RelativeHumidity, SurfacePressure, WindSpeed10m, and AlbedoDaily.

TOOLS

Existing solar panels for photovoltaic output data

Existing weather sensors used to generate publicly available weather data

Python to develop machine learning algorithms for the predictive models

Cloud services such as Amazon Web Services (AWS) to facilitate code collaboration, computing, data infrastructure and management

MODELING APPROACHES

	ML algorithms	How they work
	Polynomial regression	Models the relationship between weather variables as a continuous function.
	Random forests	Functions by constructing a tree-like model of decisions based on weather variables and can be used to identify the most important weather variables affecting solar energy output.
	LightGBM	A gradient boosting framework that uses decision trees as base learners. It can efficiently handle large-scale datasets and provide accurate predictions by iteratively learning from the mistakes of previous models.
	Extra Trees	An ensemble learning method that leverages decision trees. By training multiple decision trees on random subsets of the training data and randomly selecting features at each split, the model can make predictions based on the average or majority vote of the trees.
	CNNs (Convolutional Neural Network)	Particularly effective for capturing complex patterns and relationships in large amounts of data and can be trained on historical weather and solar energy data to make accurate predictions.
	MLP (Multilayer Perceptron)	An artificial neural network composed of multiple layers of interconnected nodes or neurons. The model can learn complex patterns and relationships in the data to make predictions.
	Bagging	Used to combine the predictions of multiple machine learning algorithms. This method works by training multiple models on different subsets of the data or using different algorithms, and then combining their predictions to produce a final prediction.
	AutoML Tool	AutoML tools were used to evaluate different models and automate tuning to compare the results of different models.

KEY FINDINGS

RESULTS

Model	RMSE (Train Data)	R-squared (Train Data)	RMSE (Test Data)	R-squared (Test Data)	RMSE (Full Dataset)	R-squared (Full Dataset)
Random Forest	1.543564972	0.953101141	4.048938569	0.675226906	2.27702766	0.897811926
LightGBM	2.810137444	0.844558289	3.95995671	0.689344859	3.074695008	0.813676477
Extra Trees w/ Hyperparameter Tuning	0.824039535	0.986633782	3.971446667	0.68753949	1.922943226	0.927121987
Bagging with Extra Trees	1.657993622	0.945889923	3.990595599	0.684519074	2.320371721	0.893884523

An evaluation was conducted on the various models to forecast solar output using weather data based on RMSE and R^2 values (see Exhibit D for full results). The results indicate that several models show promising performance in predicting solar output based on weather data. Notably, the LightGBM model stands out as the best-performing model, as it achieves the highest R^2 score on the test data among all the evaluated models. LightGBM demonstrates the highest predictive capabilities out of all the models tested with an R^2 score of 0.6893 on the test data. It outperforms other models, including Random Forest, Extra Trees, and Bagging with Extra Trees, in terms of the R^2 score on the test data. While other models, such as Random Forest, Extra Trees, and Bagging with Extra Trees, have remarkably high R^2 values for the training data, LightGBM emerges as the most accurate model when evaluated based on the test data. It achieves an RMSE value of 3.9599 on the test data, indicating its ability to make relatively precise predictions of solar output. In contrast, models like Polynomial Regression, CNN, and MLP demonstrate subpar performance on the test data, with lower R^2 scores and higher RMSE values. These models are less accurate in predicting solar output based on weather data compared to LightGBM. As such, our analysis has determined that the LightGBM model is the top-performing model for predicting solar output, as it demonstrates the highest R^2 score on the test data among all evaluated models. Its strong fit and accuracy make it a reliable choice for forecasting solar output using weather data. A scatter plot of LightGBM predicted values versus actual values from the test data (see Exhibit E) reveals a noticeable diagonal trend, indicating a general positive relationship between the predicted and actual values. While the points are somewhat spread out, they follow a loose pattern around the diagonal line. In the upper left portion of the plot, some of the predictions are higher than the corresponding actual values, resulting in a slightly greater spread. Overall, the scatter plot suggests a reasonable alignment between the predicted and actual values, with some variability in the predictions but still maintaining a general trend in the right direction. Most importantly, the LightGBM model showcased superior performance compared to all models used in previous research by Pasion et al. Their highest achieved R^2 value on test data was 0.687 from a stacked ensemble model employing a cross-validation method, while the R^2 value of the LightGBM model surpassed that benchmark.

The feature importance analysis of the LightGBM model reveals that several weather features significantly contribute to its predictive performance (see Exhibit F). Among these important features are AmbientTemp, Humidity, WindDirection10m, AirTemp, and Azimuth, which exhibit important values of 1724, 1471, 1294, 1252, and 1211, respectively. These features play a crucial role in determining the solar output predictions made by the LightGBM model. This is consistent with the correlation analysis as most of these were shown to have moderate to strong correlations with PolyPwr. The analysis of feature importance and correlation analysis collectively emphasize the significance of these weather features in accurately predicting solar output using the LightGBM model.

In analyzing the results for the Random Forest, Extra Trees, and LightGBM models by season, we found that different models perform better depending on the season (see Exhibit G). The best model for the Spring season based on the test data metrics is LightGBM, with an R^2 score of 0.5839 and an RMSE of 4.8543. The best model for the Summer season based on the test data metrics is Extra Trees, with an R^2 score of 0.6202 and an RMSE of 3.9926. The best model for the Fall season based on the test data metrics is LightGBM, with an R^2 score of 0.6819 and an RMSE of 3.4889. The best model for the Winter season based on the test data metrics is Extra Trees, with an R^2 score of 0.7901 and an RMSE of 2.7839. These findings suggest that the choice of the best model may depend on the specific season and its associated weather patterns.

When looking at the results for the Random Forest, Extra Trees, and LightGBM models by location, we found that the performance of the models varied across different locations (see Exhibit H). The Extra Trees model consistently achieved the highest R^2 scores and lowest RMSE values across most locations on the test data, indicating superior performance. However, in some locations, such as Camp Murray and March AFB, the LightGBM model also performed well. The Random Forest model generally showed lower performance compared to the other two models in most locations. Additionally, there were significant variances observed in the model performances between locations. Kahului exhibited the lowest test R^2 score with a value of 0.41 using an Extra Trees model. The highest test R^2 value was two times the lowest R^2 value at 0.82 and was observed at the Travis location also using an Extra Trees model.

The final experiment used a classification model instead of a regression model. To estimate the required power accurately, we employed a range that would fulfill our main requirement. The classification model divided target results into increments of 4 (0-3, 4-7, 8-11, etc.). Using AutoML, we achieved a Weighted AUC score of 0.81 for the test data. The XGBoost Classifier emerged as the top-performing model in this experiment. Similar to Random Forest, XGBoost classifier assigns a score to each decision within the leaves. Interestingly, Pressure emerged as the most important feature for classification problems, whereas Ambient Temperature remained significant across all the other experiments.

RECOMMENDATIONS

Based on the evaluations of the models on the entire test data set, we recommend using a LightGBM model since that was shown to have the highest R-squared and lowest RMSE. However, for improved predictive accuracy, we recommend utilizing location-based models as there was significant variability in predictive accuracy depending on location. Additionally, we recommend utilizing different modeling approaches depending on season as certain models performed better for specific seasons. The LightGBM model had the best performance for the Spring and Fall seasons while the Extra Trees model had the best performance for the Summer and Winter seasons.

IMPLEMENTATION

For this “Phase 1” model, we chose to deploy the model to a REST API and believe this approach could be replicated in subsequent iterations of the project. The model was saved using Python’s Pickle package, which saves the model parameters and weights to a custom.pkl object. To deploy, we chose to use Amazon Web Services to host the REST API through their Lambda serverless computing service. We containerized the model along with the code we used to train the model, the source data, an “api.py” python file, a “requirements.txt”, and a serverless.yml file into a github repository that could be cloned into our cloud computing instance. The api.py file contains the source code for the REST API to load the saved model, process model input data from a JSON POST call, run predictions, then return the predictions from the model to the requester using a PUT call. The requirements.txt defines the packages necessary to run the api.py service, which containerizes the app allowing for deployment on any computing service. The serverless.yml file defines the configuration for the serverless environment by defining the name of the service, the runtime, the methods, and other parameters.

We chose to deploy the model via a real-time processing API rather than a batch processing API as we felt it would better serve energy utilities to predict solar output for a given area in order to factor those numbers into their energy mix decisions for the day. A real time API could be used as an input to overall energy supply models and updated continuously or at set time increments. A second batch processing API could be created for energy supply forecasting models by feeding in 10-day weather forecasts once a day, so that utilities could plan in advance for peaks and valleys in solar energy supply.

Eventually, it may make sense to pursue a web app or dashboard that uses the REST API predictions in various visuals. If in future iterations in the project there are multiple models for each city, zip code, or neighborhood, projected energy production for each of those locations/models could be visualized next to each other, so that utilities could be more precise in routing energy to areas that are at risk of not meeting demand. However, further refinement of the model to increase predictive ability and location specificity would be needed before a web application would be worth the development effort.

CONCLUSION

Solar Tech Solutions provides competitive accuracy in predicting solar power output. By applying various models to the weather features training data, we were able to outperform existing research by producing an R^2 value on test data that exceeds current results using LightGBM, which was the best performing model out of the various models trained. Key weather features were identified by performing feature importance analysis on the data, which confirmed our findings in our correlation analysis in the preliminary EDA. We furthered our analysis by modeling against key variables, such as location and season, and found that certain locations and seasons performed better using Extra Trees. Finally, using a classification model we identified XG Boost Classifier to perform the best using Pressure as the main feature. Having identified the models that provide the most accurate predictions by appropriate seasons and locations, we can apply these results to a REST API and ultimately a web app or dashboard that is user friendly and provides accurate predictions. Clients and utility companies can then use this information to make strategic decisions in how to optimally allocate and utilize the solar energy absorbed.

MEET THE TEAM

Shahar Journo
Project Manager & Solution Architect

Andrew Gray
Data Scientist

Scott Jue
Data Analyst

Johanna Sheu
Data Analyst

APPENDIX

EXHIBIT A: AZURE AUTOML MODELS CHOSEN FOR REGRESSION MODELS

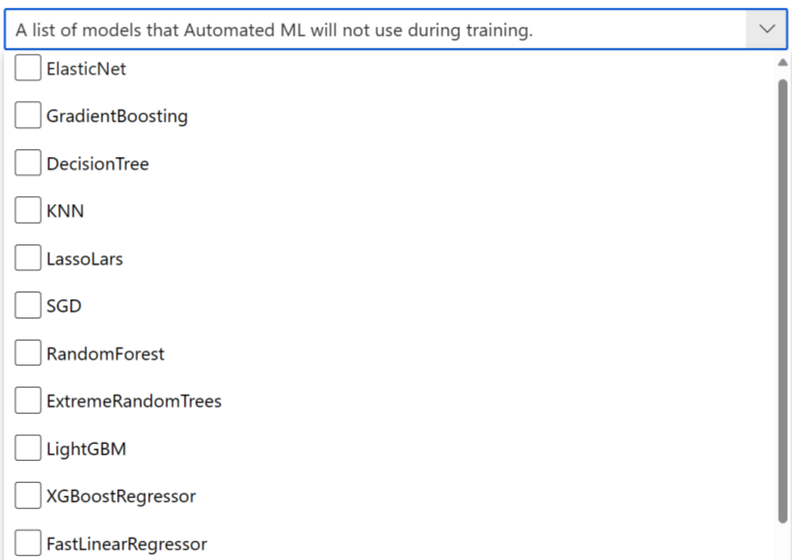


EXHIBIT B: AZURE AUTOML MODELS CHOSEN FOR CLASSIFICATION MODEL

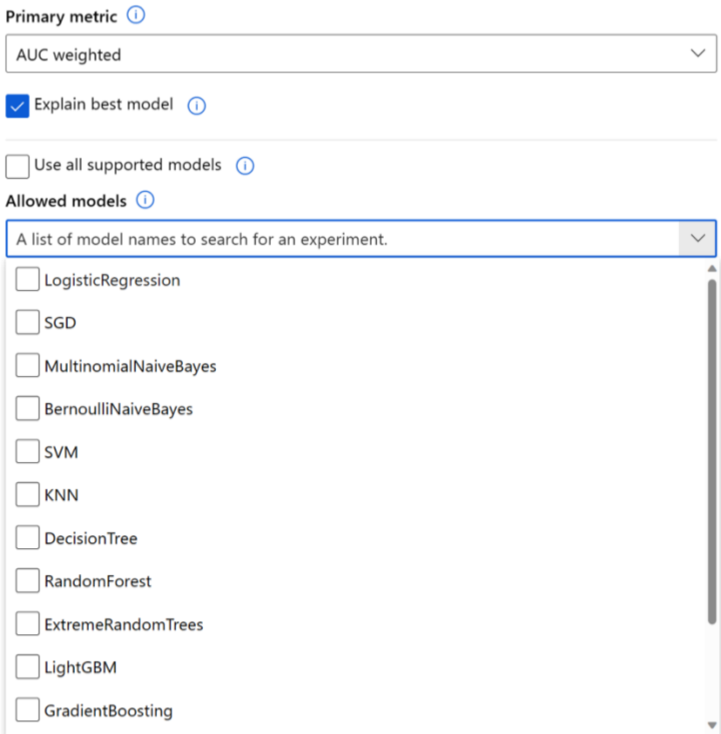


EXHIBIT C: EXPERIMENT 4 CLASSIFICATION MODEL – TOP 4 FEATURES FOR OVERALL MODELS

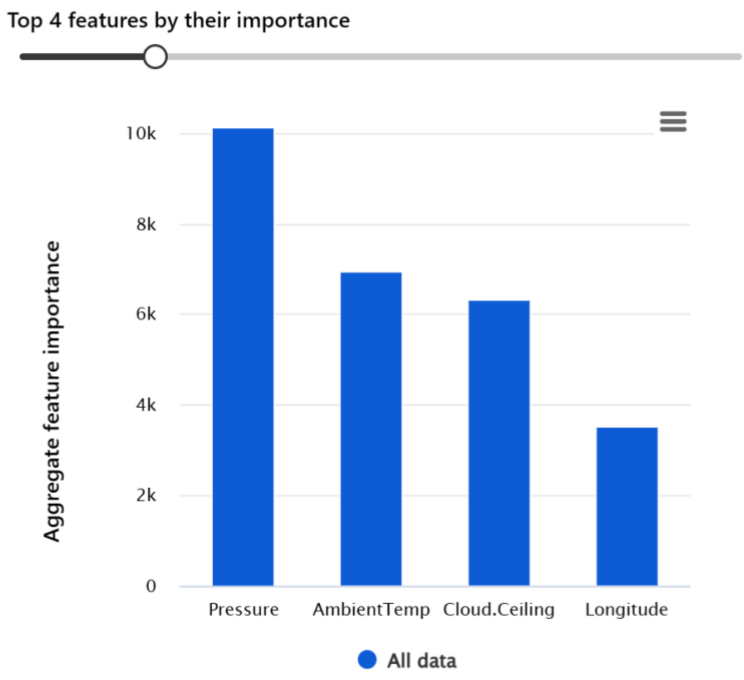


EXHIBIT D: RESULTS OF OVERALL MODELS

Model	RMSE (Train Data)	R-squared (Train Data)	RMSE (Test Data)	R-squared (Test Data)	RMSE (Full Dataset)	R-squared (Full Dataset)
Random Forest	1.543564972	0.953101141	4.048938569	0.675226906	2.27702766	0.897811926
Random Forest w/ Hyperparameters Tuning	2.426062151	0.884144598	4.033884476	0.677637451	2.82189065	0.843056273
Random Forest w/ Automated Feature Selection	1.574279364	0.951216152	4.074492202	0.671114558	2.302820391	0.89548377
LightGBM	2.810137444	0.844558289	3.95995671	0.689344859	3.074695008	0.813676477
Extra Trees	2.06566E-14	1	3.98565804	0.68529928	1.782440463	0.937382797
Extra Trees w/ Hyperparameter Tuning	0.824039535	0.986633782	3.971446667	0.68753949	1.922943226	0.927121987
Bagging with Extra Trees	1.657993622	0.945889923	3.990595599	0.684519074	2.320371721	0.893884523
Polynomial Regression	3.727786242	0.726464029	4.373561379	0.621062034	3.865581462	0.705494601
CNN	5.030153312	0.501947771	4.979557774	0.508776405	5.020074996	0.503311606
MLP	4.939944758	0.51965129	4.895804254	0.525161681	5.020074996	0.503311606

EXHIBIT E: PREDICT VS ACTUAL VALUES – LIGHTGBM (TEST DATA)

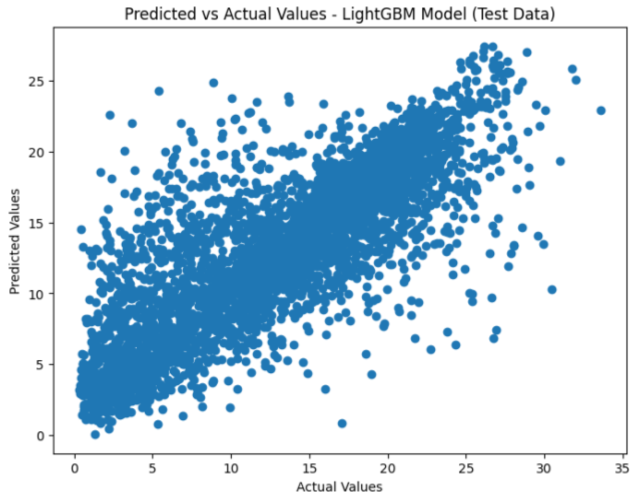


EXHIBIT F: LIGHTGBM FEATURE IMPORTANCE

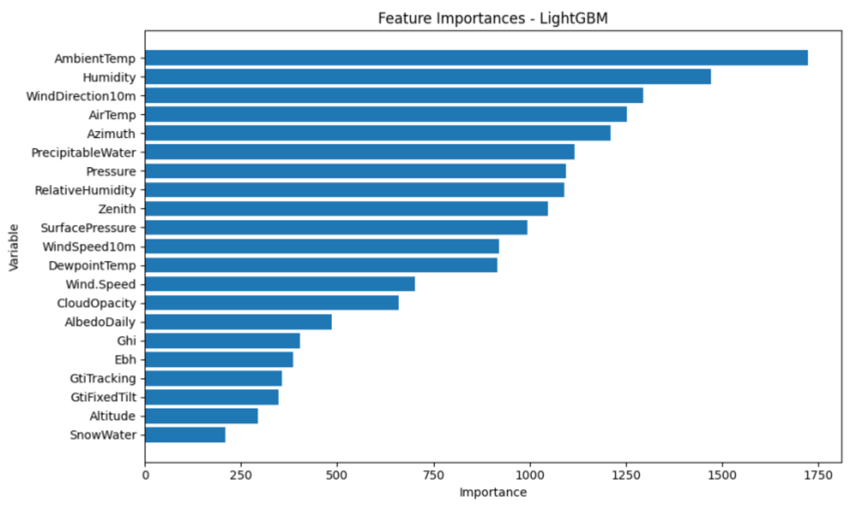


EXHIBIT G: RESULTS OF MODELS BY SEASON

Random Forest by Season:

Season	R^2 Score (Train Data)	RMSE (Train Data)	R^2 Score (Test Data)	RMSE (Test Data)	R^2 Score (Full Data)	RMSE (Full Data)
Fall	0.949174955	1.378851699	0.678026693	3.509986034	0.893518682	2.000738231
Spring	0.937504458	1.902210255	0.582885156	4.860037922	0.866967338	2.769530924
Summer	0.941389831	1.577443197	0.611407796	4.038442871	0.876344194	2.288664469
Winter	0.964737645	1.121549007	0.755815609	3.003005631	0.922474001	1.668976241

Extra Trees by Season:

Season	R^2 Score (Train Data)	RMSE (Train Data)	R^2 Score (Test Data)	RMSE (Test Data)	R^2 Score (Full Data)	RMSE (Full Data)
Fall	1	1.77256E-14	0.665199657	3.579220066	0.931257014	1.607562093
Spring	1	2.18861E-14	0.589555283	4.821022635	0.91833825	2.16988351
Summer	1	2.37129E-14	0.620183926	3.992579495	0.925130275	1.780852665
Winter	1	1.36857E-14	0.790147404	2.783904737	0.957538348	1.235164815

LightGBM by Season:

Season	R^2 Score (Train Data)	RMSE (Train Data)	R^2 Score (Test Data)	RMSE (Test Data)	R^2 Score (Full Data)	RMSE (Full Data)
Fall	0.872766744	2.181619445	0.681880977	3.488914079	0.833616443	2.500971621
Spring	0.826295209	3.171318768	0.583862074	4.854343279	0.778109967	3.576810861
Summer	0.794565523	2.953275121	0.604781623	4.07272857	0.75715777	3.207280754
Winter	0.923906595	1.64754053	0.769829347	2.915561263	0.892751518	1.963007884

EXHIBIT H: RESULTS OF MODELS BY LOCATION

Random Forest by Location:

Location	R^2 Score (Train Data)	RMSE (Train Data)	R^2 Score (Test Data)	RMSE (Test Data)	R^2 Score (Full Data)	RMSE (Full Data)
Camp Murray	0.965340618	1.293266681	0.728656589	3.429332925	0.923061256	1.907705343
Grissom	0.946018753	1.578942742	0.678206139	3.747400321	0.893787321	2.20325079
Hill Weber	0.954167246	1.459306555	0.68554031	3.823460566	0.902425961	2.129361448
JDMT	0.934549861	1.918287008	0.528914211	5.070545304	0.856633747	2.8309222
Kahului	0.918934642	2.029834668	0.417385148	5.674197712	0.803107648	3.194824891
MNANG	0.950937742	1.709978366	0.693452494	4.221721804	0.901628645	2.417954572
Malmstrom	0.95054073	1.561508701	0.687360497	4.027675684	0.894563851	2.29635008
March AFB	0.948098262	1.159838489	0.655484677	3.049027027	0.889103981	1.702979182
Offutt	0.945038943	1.904224278	0.577223371	5.221781469	0.870003067	2.921768128
Peterson	0.942685695	1.622625401	0.63339478	4.118003174	0.881790738	2.332001914
Travis	0.968343028	1.18840701	0.804747572	2.934293343	0.934164069	1.711717176
USAFA	0.936070217	1.481054087	0.506529594	4.029104935	0.855472226	2.218532405

Extra Trees by Location:

Location	R^2 Score (Train Data)	RMSE (Train Data)	R^2 Score (Test Data)	RMSE (Test Data)	R^2 Score (Full Data)	RMSE (Full Data)
Camp Murray	1	1.85031E-14	0.74598261	3.318040888	0.954621458	1.465089036
Grissom	1	1.68475E-14	0.674558126	3.768581647	0.936471739	1.703959375
Hill Weber	1	2.23474E-14	0.698443574	3.744194652	0.941915697	1.642901915
JDMT	1	2.58887E-14	0.514300287	5.148593322	0.906704576	2.283676786
Kahului	1	2.18267E-14	0.409886618	5.710595783	0.863623948	2.658898092
MNANG	1	1.85165E-14	0.701806568	4.163798987	0.942779886	1.844115717
Malmstrom	1	1.70137E-14	0.713069336	3.858522604	0.938775612	1.749869072
March AFB	1	2.09518E-14	0.646348497	3.089190972	0.92864039	1.366084605
Offutt	1	2.16866E-14	0.593384534	5.121004458	0.917047954	2.333956431
Peterson	1	1.85108E-14	0.619415783	4.195780319	0.925061361	1.856761189
Travis	1	2.1453E-14	0.819897116	2.818159966	0.962372248	1.294060725
USAFA	1	2.0556E-14	0.51860483	3.97950346	0.90963344	1.754260141

LightGBM by Location:

Location	R^2 Score (Train Data)	RMSE (Train Data)	R^2 Score (Test Data)	RMSE (Test Data)	R^2 Score (Full Data)	RMSE (Full Data)
Camp Murray	0.975738652	1.082019651	0.723262932	3.463248674	0.930637401	1.811346051
Grissom	0.96211502	1.322752092	0.690303305	3.676287758	0.909088852	2.038376126
Hill Weber	0.944465616	1.606347621	0.694843078	3.766480625	0.896384928	2.194288425
JDMT	0.942779957	1.79362806	0.508715538	5.178108906	0.859403085	2.803447083
Kahului	0.951020304	1.577796734	0.432244922	5.601369241	0.831180001	2.958316143
MNANG	0.972601011	1.277862476	0.726661805	3.986491839	0.925463518	2.10473904
Malmstrom	0.964828826	1.316782344	0.694872722	3.978992172	0.907354215	2.152564448
March AFB	0.939430133	1.252953275	0.673401225	2.968686099	0.8858083	1.728099051
Offutt	0.971714041	1.366079614	0.607343121	5.032338104	0.897380529	2.595936582
Peterson	0.931662442	1.771807281	0.611827313	4.2374037	0.868692387	2.457808845
Travis	0.954461747	1.425340758	0.808554223	2.905548992	0.923978215	1.839371903
USAFA	0.898612616	1.865139264	0.47361494	4.161307407	0.818887652	2.483499919

CODE

Our team used Colab to integrate the team's coding efforts.

Link to the repository may be found here https://colab.research.google.com/drive/1Pk5K-GKnaAz_XFTU92qjRNfYTewH6LUH

LITERATURE REVIEW

Dabaghzadeh, N., and M. Eslami. 2019. "Temperature distribution in a photovoltaic module at various mounting and wind conditions: a complete CFD modeling." Journal of Renewable and Sustainable Energy 11, no. 5. Article 053503.

Doddy Clark, Eadaoin, and Conor Sweeney. 2021. "Solar energy and weather" Royal Meteorological Society. https://doi.org/10.1002/wea.4124.

Huld, Thomas, Gabi Friesen, Artur Skoczek, Robert P. Kenny, Tony Sample, Michael Field, and Ewan D. Dunlop. 2011. "A power-rating model for crystalline silicon PV modules" Solar Energy Materials and Solar Cells Vol. 95, Iss. 12. https://doi.org/10.1016/j.solmat.2011.07.026

Kim, Alex, and Dane Stocks. 2019. Rep. "Predicting Solar Power Generation from Weather Data" Palo Alto, CA. http://cs229.stanford.edu/proj2019spr/report/91.pdf.

Omran, Shaimaa, Murat Dilek, and Robert Broadwater. 2020. "Global Horizontal Irradiance and Meteorological Parameters Analysis for Modeling Effects of Cloud Motion on PV Generation" International Journal of Renewable Energy Research Vol. 10, No. 1. https://doi.org/10.20508/ijrer.v10i1.10507.g7868.

Pasion, Christil, Torrey Wagner, Clay Koschnick, Steven Schuldt, Jada Williams, and Kevin Hallinan. 2020. "Machine Learning Modeling of Horizontal Photovoltaics Using Weather and Location Data" Energies 13, no. 10: 2570. https://doi.org/10.3390/en13102570.

Prabhakar, Akash. 2021. "Solar Power forecasting using Machine Learning and Deep Learning" Medium. https://medium.com/@akashprabhakar427/solar-power-forecasting-using-machine-learning-and-deep-learning-61d6292693de

You, Zhiyuan, Lei Cui, Yujun Shen, Kai Yang, Xin Lu, Yu Zheng, and Xiny Le. 2022. "A unified model for multi-class anomaly detection" arXiv. https://doi.org/10.48550/arXiv.2206.03687.

Zazoum, Bouchaib. 2022. "Solar Photovoltaic Power Prediction Using Different Machine Learning Methods" Energy Reports 8: 19–25. https://doi.org/10.1016/j.egyr.2021.11.183.