TECHNICAL UNIVERSITY OF CRETE
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Στατιστική Μοντελοποίηση και Αναγνώριση Προτύπων

Μεταπτυχιακό Project

Ευστάθεια ηλεκτρικού δικτύου με Έξυπνα Δίκτυα

Ιάσων Χρήστος Καλαϊτζάκης
AM: 2020039003

June 2021

Introduction

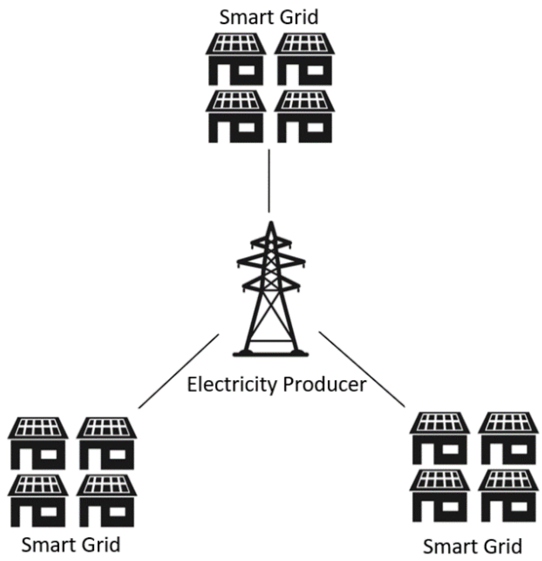
Since the introduction of “smart grids” to the general grid, the bidirectional flow of power has created new stability issues to a grid that once was one-directional and easy to manage. A smart grid is a consumer node that can also produce power, hence the bidirectionality.

Apart from the bidirectionality, however, another issue is the decentralization of power management. Since the power producers are no longer centralized (usually government owned) power plants, there needs to be a properly controlled producer/consumer environment to ensure the economic stability of power transactions.

Combining the above two facts, one can see how grid stability and its detection is of paramount importance to the health of the decentralized grid. The project’s goal is to assess grid stability depending on certain parameters of the grid’s members.

System Description

The examined system is a four-star node, with 3 smart grid “prosumers” in star formation and a power producer at the center.



The stability of any power system is measured through its frequency. Frequency deviations of more than 0.5Hz can cause grid instability issues, and deviations of more than 1Hz can force the whole grid into a blackout, since generators would have to be shut off.

The basic principle of each node’s operation is described by the following differential equation, provided by [1]:

$$\frac{d^2\theta_j}{dt^2} = P_j - a_j\frac{d\theta_j}{dt} + \sum_{k=1}^{N} K_{jk}\sin\left(\theta_k - \theta_j\right) - \frac{\gamma_j}{\mathrm{T}_j}\left(\theta_j\left(t-\tau_j\right) - \theta_j\left(t-\tau_j-\mathrm{T}_j\right)\right)$$

The basic parameters that are important to our classification project are the following:

θ_j → rotor angle (proportionate to the grid's frequency)

P_j → mechanical energy that is produced/consumed

τ_j → reaction time of each node to a fluctuation in the energy's price

γ_j → price elasticity, an economical metric that describes how much the energy demand/supply will change depending on a change in energy price

Dataset

The dataset's source is [2]. However, the creator of the aforementioned dataset sourced it from [3] and enhanced it, increasing the number of samples from 10000 to 60000. The original dataset and the enriched dataset are both simulated data, which resulted from the research work [1] on Decentral Smart Grid Control.

The dataset contains the following 12 features:

- *τ [1-4]*→ Reaction time of the producer (τ1) and the consumers (τ2, τ3, τ4)
- *ρ [1-4]*→Produced (ρ1) and consumed (ρ2, ρ3, ρ4) power
- *γ [1-4]*→ Price elasticity of the producer (γ1) and the consumers (γ2, γ3, γ4)

The 12 features will be analyzed later, for the purpose of dropping the ones that do not assist our classification model.

The dependent variables are the following:

- *stabf* → The real part of the characteristic differential equation's largest root

- *stab* → Binary variable (0 or 1) depending on the sign of stabf

As is common knowledge in control theory, a root with positive real part indicates an unstable system. Therefore, stab=0 (unstable) for positive stabf and stab=1 (stable) for negative stabf.

The following table is an example of 4 random samples from the dataset.

	Predictive Features												Dependent Variables	
	τ1 (s)	τ2 (s)	τ3 (s)	τ4 (s)	p1 (kW)	p2 (kW)	p3 (kW)	p4 (kW)	γ1	γ2	γ3	γ4	stabf	stab
Samples	2.95906	3.079885	8.381025	9.780754	3.763085	-0.7826	-1.25739	-1.72309	0.650456	0.859578	0.887445	0.958034	0.055347	0 (unstable)
	9.304097	4.902524	3.047541	1.369357	5.067812	-1.94006	-1.87274	-1.25501	0.413441	0.862414	0.562139	0.78176	-0.00596	1 (stable)
	8.971707	8.848428	3.046479	1.214518	3.405158	-1.20746	-1.27721	-0.92049	0.163041	0.766689	0.839444	0.109853	0.003471	0 (unstable)
	6.999209	9.109247	3.784066	4.267788	4.429669	-1.85714	-0.6704	-1.90213	0.261793	0.07793	0.542884	0.469931	-0.01738	1 (stable)

The reaction times range from 0 to 10 seconds, and the power values range from 0 to 6kW for the producer (p1) and -2 to 0kW for the consumers (p2, p3, p4). Price elasticity is a dimensionless variable ranging from 0 (inelastic response, does not respond to price fluctuations) to 1 (elastic response, responds well to price fluctuations).

As far as the dependent variables are concerned, the samples with a positive stabf result in instability, and negative stabf in stability, as expected from what was mentioned before.

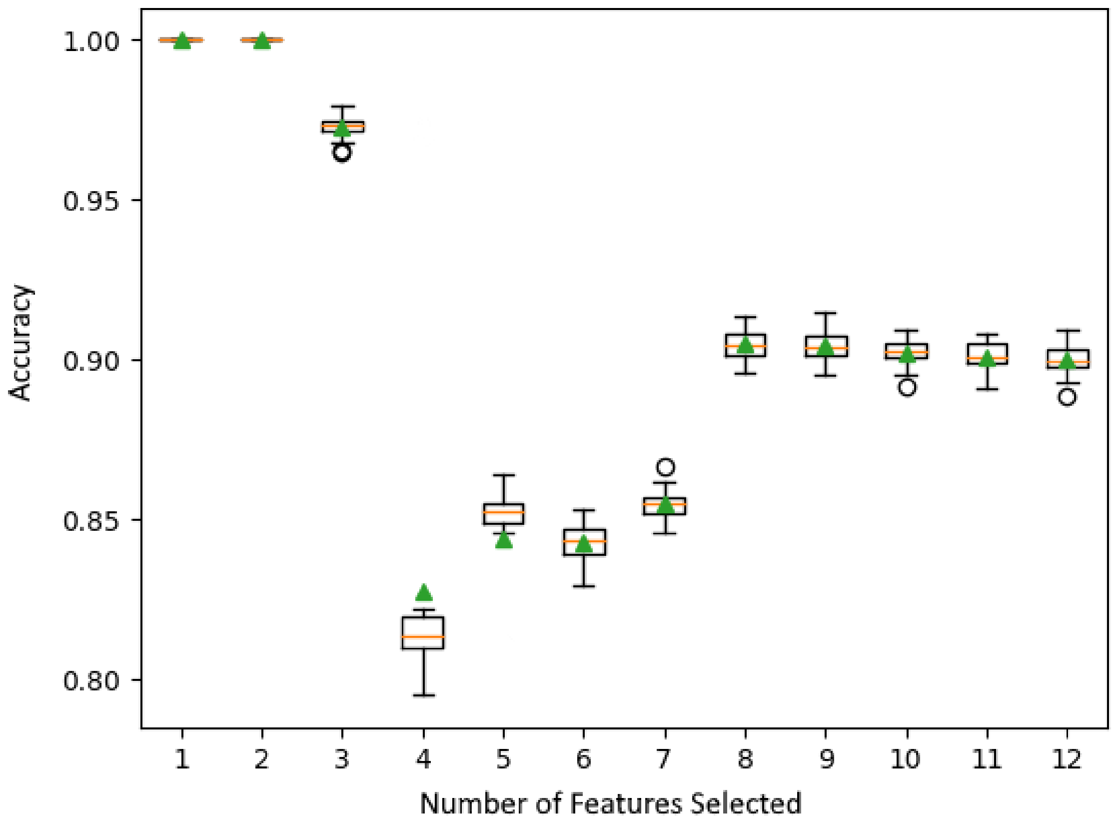
Feature extraction and analysis

The first step to evaluating our features is the creation of the correlation map. Through the correlation map, we can see how much each feature affects the final result of stability.

The map's most important values are of the last row, where we can see each feature's correlation with stab, our dependent variable. With 0 stab meaning unstable grid, and 1 stab meaning stable grid, a positive correlation with it means that as the feature increases in value, so does the stability.

The four power features have very little correlation with the stability, which means that they may need to be dropped for more accurate results. Elasticity and response time have positive correlation, which means a stable grid should display high elasticity and response time.

To further analyze our suspicion that the four power features must be dropped, Recursive Feature Elimination (RFE) is applied on the dataset. RFE is a feature selection algorithm that measures the possible accuracy of the classification or regression model depending on the number of features selected. The algorithm can output the optimal number of selected features, as well as rank them from most important to least important.



It would appear that the algorithm shows that selecting 1-3 features is the optimal solution for classification accuracy. However, this could result in an overfitted, non-generalized classifier, and thus this specific result is ignored. According to the next number of features with the highest accuracy, the final decision would seem to be 8 features.

The following table displays the ranking of features by RFE:

τ1	τ2	τ3	τ4
Selected, Rank: 1	Selected, Rank: 2	Selected, Rank: 3	Selected, Rank: 4
ρ1	ρ2	ρ3	ρ4
Dropped, Rank: 8	Dropped, Rank: 10	Dropped, Rank: 9	Dropped, Rank: 11
γ1	γ2	γ3	γ4
Selected, Rank: 1	Selected, Rank: 6	Selected, Rank: 5	Selected, Rank: 7

Confirming our earlier suspicions, the four power features are the least important, and are therefore dropped. Another interesting thing to note is the ranking of the producer, with the producer's features outranking the consumers' features. That is to be expected, as the power producer node is by far the most important one in grid stability.

Support Vector Machine

Moving on to classification, our first classifier is a Support Vector Machine (SVM).

The first important decision regarding an SVM is to select the kernel that is going to transform the dataset into a classifiable set of samples.

The SVM library used in python offers three different kernels, Linear, Polynomial and Radial basis function (RBF). The classification problem is executed with all three in order to compare them and select the optimal one.

Linear Kernel: $K(x,\ xi)\ =\ sum(x\ *\ xi)$

Accuracy: 81%

Polynomial Kernel: $K(x, xi) =\ 1\ +\ sum(x\ *\ xi)^{d},\ d = 3$

Accuracy: 96%

Radial Basis Function Kernel: $K(x, xi) = e^{-gamma * sum(x - xi)^2}$

Accuracy: 94%

According to the results, the polynomial kernel of the third degree is the most efficient one, and is thus selected.

The dataset is split into 31000 samples for training, 11000 samples for testing, and 18000 samples for blind testing.

The following are the results of the SVM in classifying the blind testing data:

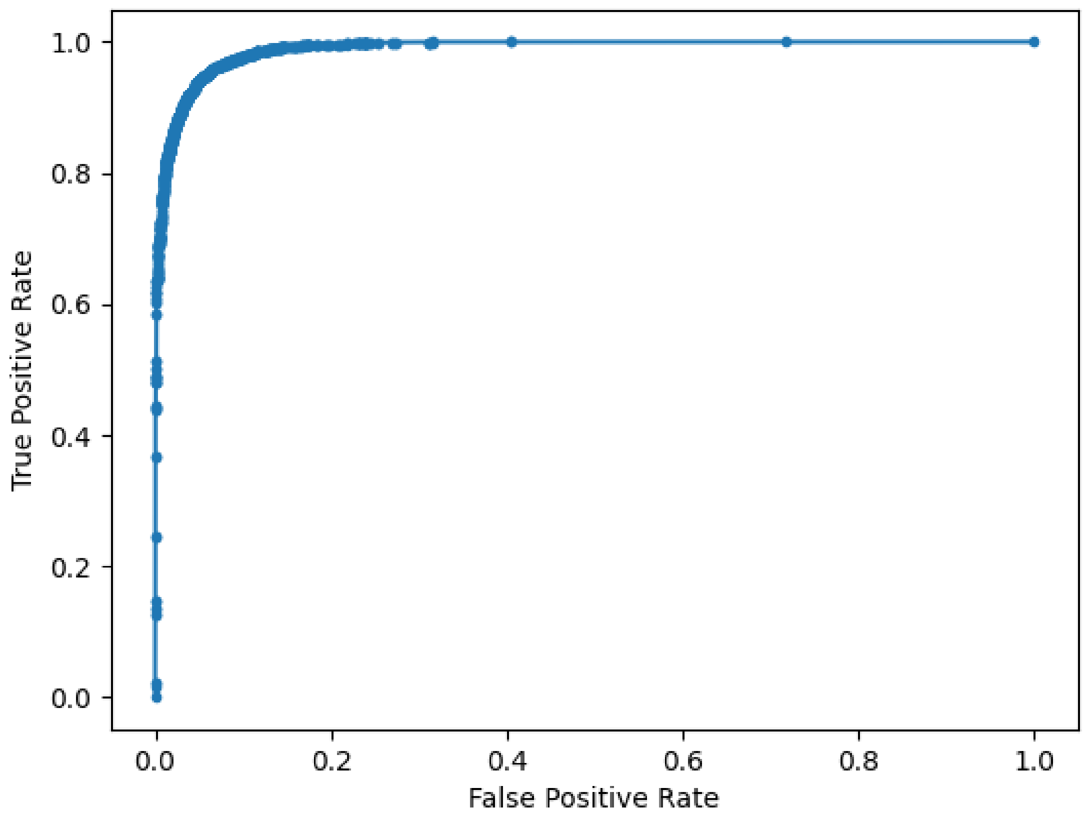
- Accuracy: 96%
- Precision: 92%
- Recall: 93%
- Sensitivity: 93%
- Specificity: 95%

and the confusion matrix:

	Predicted Negative (unstable)	Predicted Positive (stable)
Actual Negative (unstable)	**10938**	519
Actual Positive (stable)	456	**6087**

It can be concluded that the SVM exhibits satisfactory behavior in classifying our dataset, with very few of the samples being misclassified (not belonging in the diagonal of the confusion matrix).

Another metric for measuring the performance of the classifier is the Receiver Operating Characteristic (ROC) curve. The ROC curve displays the true positive rate as the decision boundary of the algorithm changes, starting from 0 to the left (always negative) to 1 to the right (always positive). The behavior of the classifier between 0 and 1 measures the quality, with the Area Under the Curve (AOC) of the ROC curve being the correct classification of the sample. Thus, a good classifier has a high AOC metric, which means it selects the correct class most of the time, despite the unfavorable decision boundary.

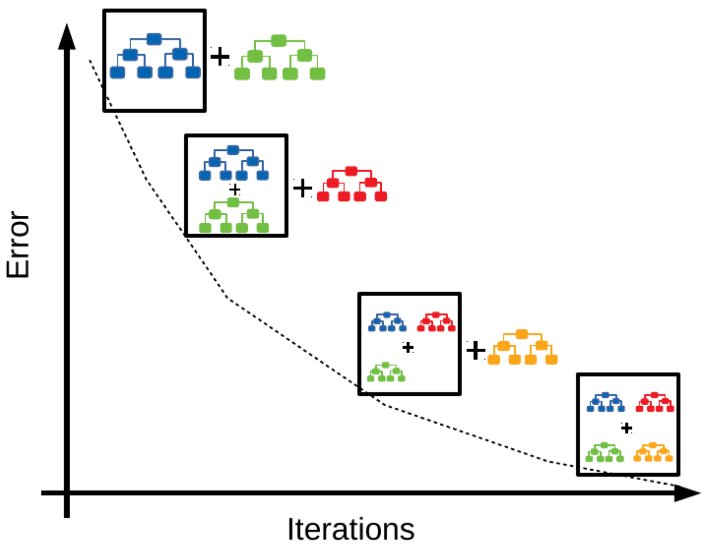


The SVM classifier exhibits a very solid ROC curve, with a negligible area above the curve. The AOC measures up to 98%, which, as mentioned before, is a great value regarding the quality of the classification.

Gradient Boosting

A more technologically recent and widely chosen algorithm for classification is that of Gradient Boosting (GB). Gradient Boosting is part of a category of classifiers known as Ensemble Algorithms, which are algorithms that combine the decisions of many classifiers in order to reach a more accurate result.

Gradient Boosting specifically employs many Binary Decision Tree classifiers, and uses majority vote to select the correct class. The boosting part of the algorithm comes from the fact that the weaker classifiers learn from the mistakes of the previous ones as iterations increase.



By the end of the iterations, the decision trees have improved to the point where their majority can correctly select the class.

The following are the results of the GB in classifying the blind testing data:

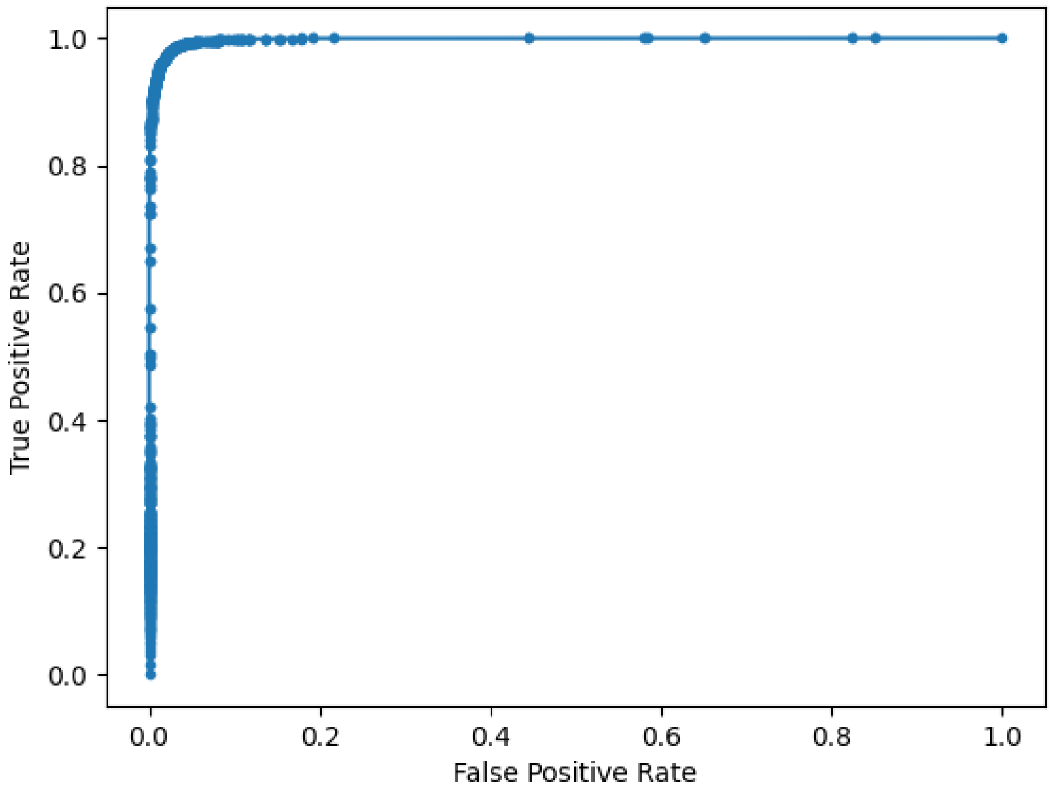
- Accuracy: 98%
- Precision: 97%
- Recall: 96%
- Sensitivity: 96%
- Specificity: 99%

and the confusion matrix:

	Predicted Negative (unstable)	Predicted Positive (stable)
Actual Negative (unstable)	**11317**	168
Actual Positive (stable)	255	**6260**

Comparing the results to the results of the SVM, it is noticeable that the GB classifier had better accuracy, resulting in almost half the misclassified samples.

Additionally, it should be noted that the GB classification had less computational complexity, as it was executed in a few seconds, compared to the SVM that needed about a minute to run. This could be a result of better multi-threaded behavior, however.



The ROC curve of the GB classifier is almost perfect, with an AOC of 99.8%.

Conclusions

To summarize, both Support Vector Machine and Gradient Boosting classification exhibited satisfactory behavior in classifying the dataset, with GB being slightly more superior.

Regarding the dataset, we can conclude that the power consumption/production itself does not affect the stability of the grid, but the most important factors are the price elasticity and the response delay.

References

[1] V. Arzamasov, K. Böhm and P. Jochem, "Towards Concise Models of Grid Stability," *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2018, pp. 1-6, doi: 10.1109/SmartGridComm.2018.8587498.

[2] https://www.kaggle.com/pcbreviglieri/smart-grid-stability, Paulo Breviglieri, “Smart Grid Stability: Augmented version of the original hosted in the UCI Machine Learning Repository.”

[3] https://archive.ics.uci.edu/ml/datasets/Electrical+Grid+Stability+Simulated+Data+#, Vadim Arzamasov, “Electrical Grid Stability Simulated Data Data Set”