Predicting Power Grid Component Outage In Response to Extreme Events

Rozhin Eskandarpour & Amin Khodaei
University of Denver
USA

Shay Bahramirad
ComEd
USA

Math Bollen
Luleå Univ of Technology
Sweden

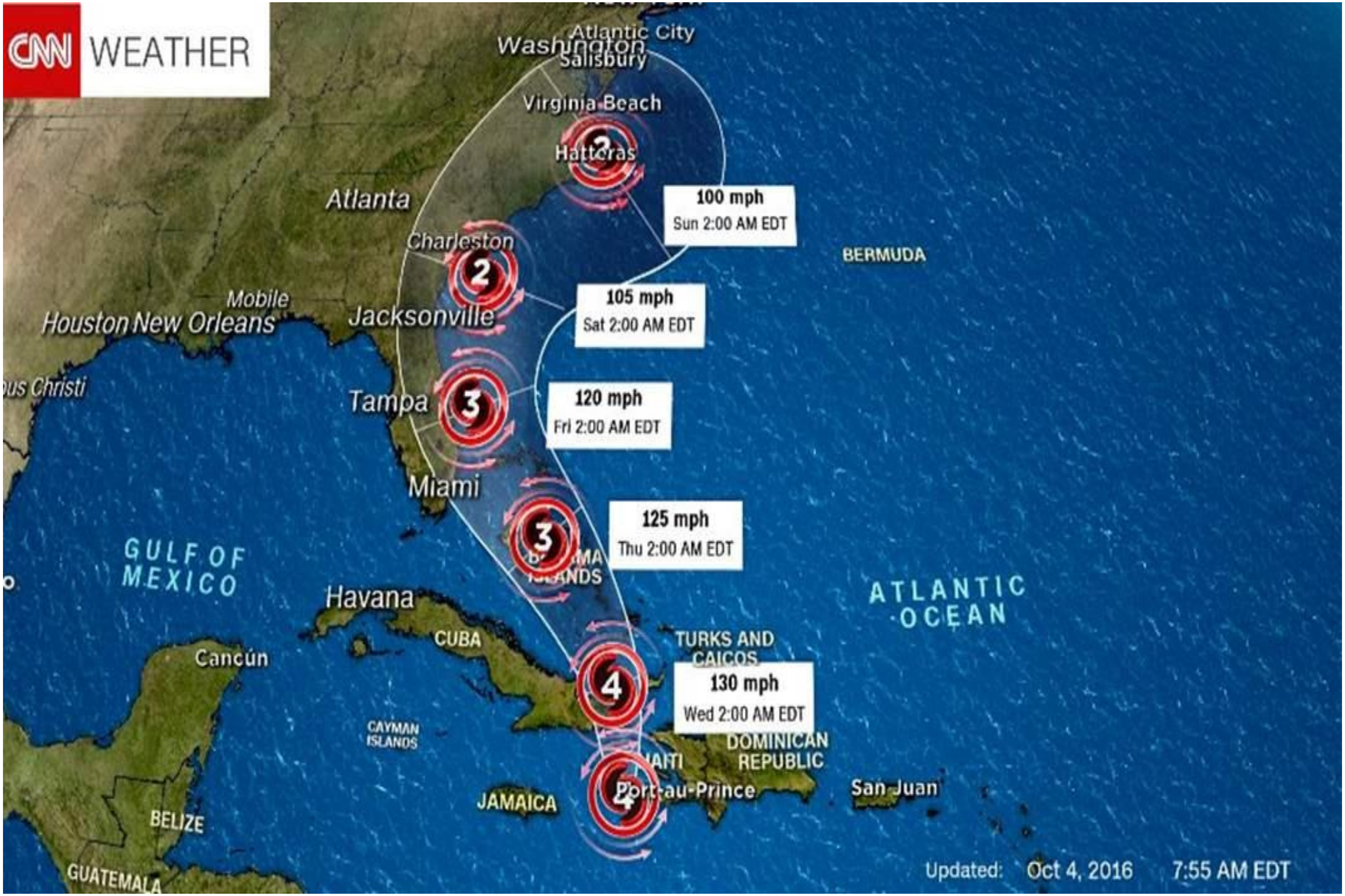
Introduction

- Extreme events, including severe weather events and natural disasters, cause significant economic, social, and physical disruptions, and result in considerable inconvenience for residents living in disaster areas due to loss of critical lifeline systems.

- The electricity infrastructure has always been significantly impacted by these extreme events as it is dispersed over a vast geographical area and hence prone to be largely affected.

- An accurate prediction of the outage of power system components in response to extreme weather events is of ultimate importance in ensuring a viable resource management, restoration, and recovery.

Hurricane Matthew



Machine Learning

Machine learning methods can be used for prediction purposes.

- Machine learning includes data-driven decision-making techniques that explore algorithms that are able to learn from, describe, and make predictions on data.

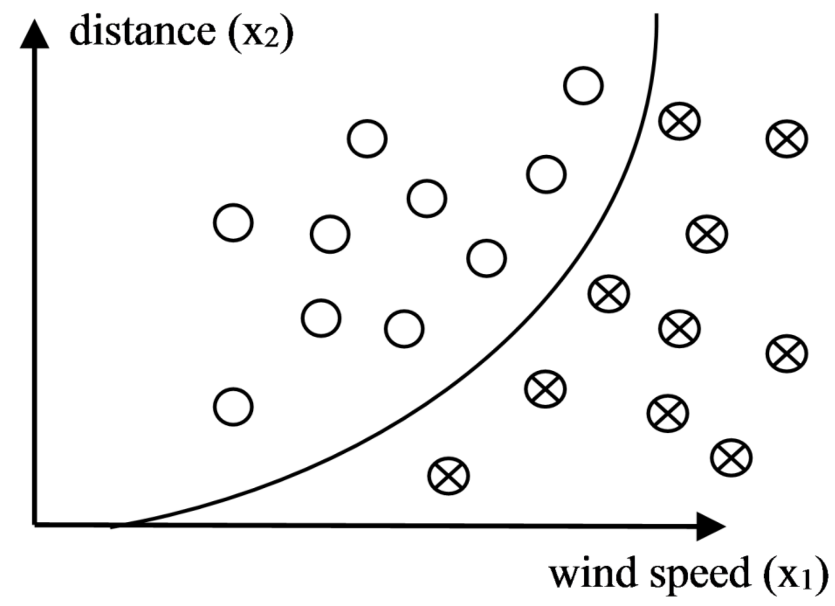
- These algorithms can make use of the available data to model the relationships between different variables of interest (regression), categorize observations based on qualitative features (classification), or combine similar patterns into homogeneous groups (clustering).

Logistic Regression

- As first step, a machine learning method, based on logistic regression, is proposed to predict the system components that can potentially fail during an expected extreme event.

- Let's say we have historical data for the path and the intensity of an extreme event.
- Further consider two states for each grid component: damaged (on outage) and operational (in service)

- We can find a decision boundary between damaged and operational components



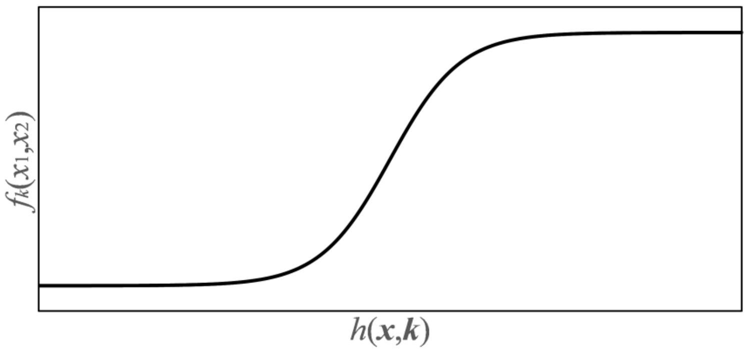
Logistic Regression Formulation

- The decision boundary is defined by a second order polynomial based on the wind speed and the distance:

$$h(x,k) = k_0 + k_1 x_1 + k_2 x_2 + k_3 x_1^2 + k_4 x_2^2 + k_5 x_1 x_2$$

- The classification function is denoted by f(x,k) and defined as a Sigmoid function, i.e.,

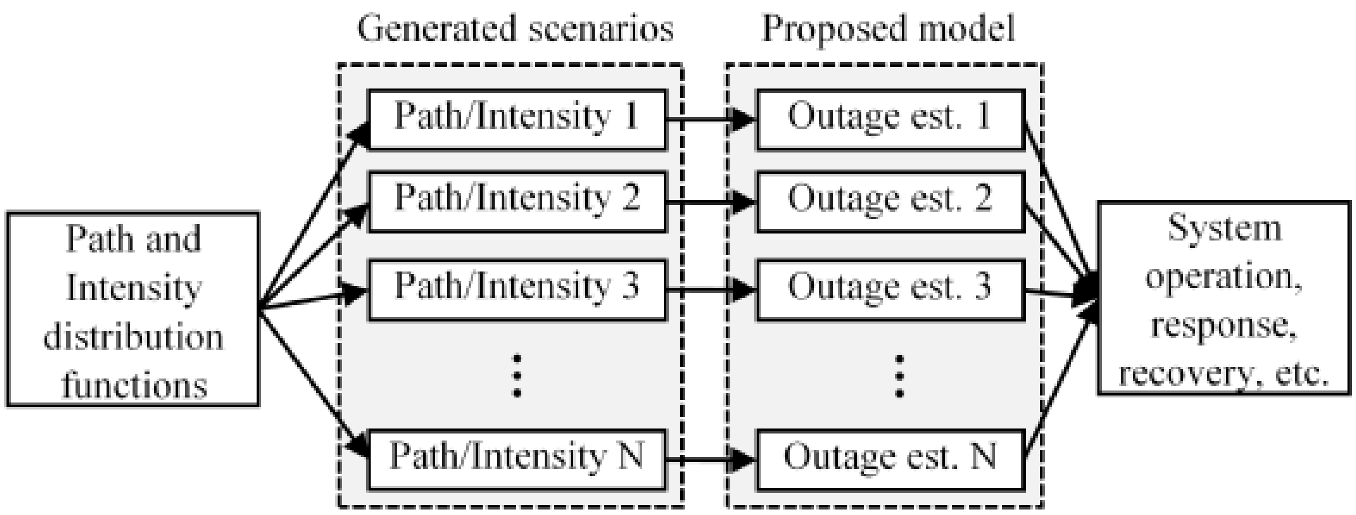
$$f(x,k) = \frac{1}{1+e^{-h(x,k)}}$$



- This function ensures that for positive values of h(x,k) a value of 1 is reached, while for its negative values, a value of 0 is reached. This function nicely classifies the data based on the obtained function.

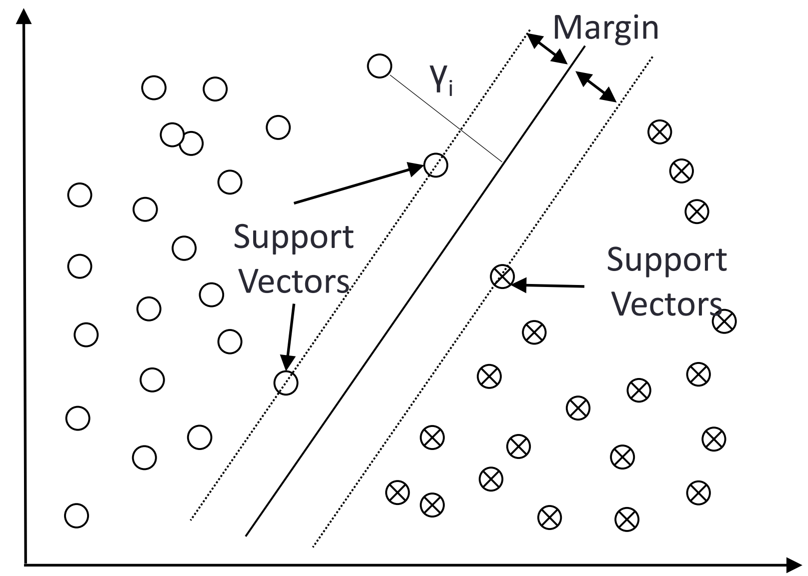
Uncertainty Consideration

- To consider uncertainty in the path and speed of the extreme event, we can generate scenarios, i.e., a scenario-based outage prediction model.



Support Vector Machine

- Support Vector Machine (SVM) is also used for classification, but it can classify the data into two classes by finding the best hyperplane that separates training examples of one class from the other class.



- The data for training is a set of points x_i ($x_i \in R_D$) along with their categories y_i ($y_i = \pm 1$), the classification task can be written as:

$$h_{w,b}(x) = g\left(w^T x + b\right). \qquad \hat{g}_i = y_i\left(w^T x_i + b\right)$$

Evaluation

- To evaluate the performance of the classifier, usually a subset of historical data is reserved as the validation/test set.

- The F_1-Score is a common and reliable measure of classification performance:

$$F_1 = \frac{2PR}{(P+R)}$$

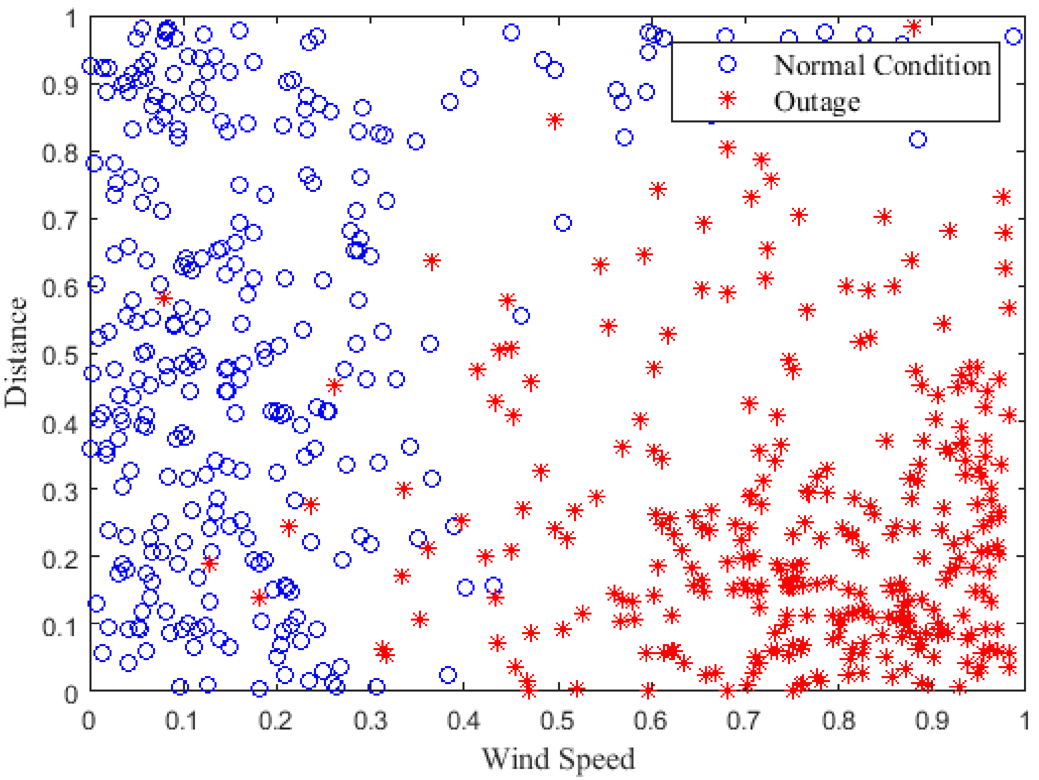
$$P = \frac{\text{number of correctly predicted outages}}{\text{total number of predicted outages}}$$

$$R = \frac{\text{number of correctly predicted outages}}{\text{total number of actual outages}}$$

Case Study

- As historical data for the past extreme events at component level are limited, we generated 300 samples of each component state following a normal distribution function with a small Gaussian noise so that the data can be distinguishable.

- The samples belong to two classes of components with high probability of failure and components that can survive the extreme event.
- The features are normalized to [0, 1] based on the maximum considered values of wind speed and distance.

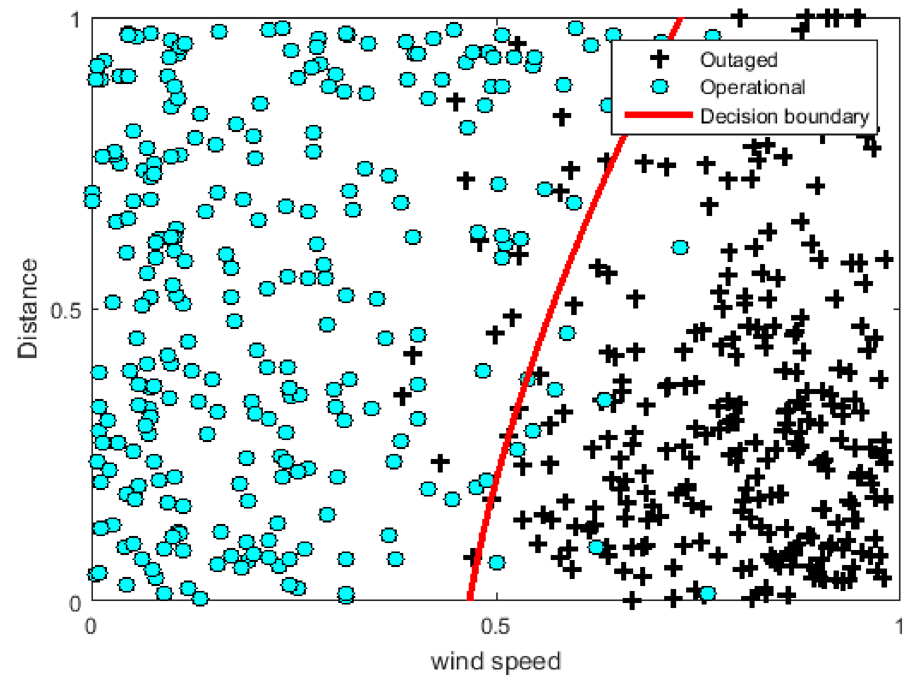


Logistic Regression Performance

- The proposed method results in the characteristics parameters of the decision boundary as: k_0 = -1.47, k_1 = 2.85, k_2 = -0.59, k_3 = 2.05, k_4 =-0.70, and k_5 =0.36.

- F_1-score = 0.90

		Predicted	
		Operational	Outage
Actual	Operational	**87.09%**	12.91%
	Outage	6.89 %	**93.11%**



SVM Performance

- Different kernels (linear, polynomial quadratic and cubic) with different range of penalty parameter (c=0.01, 0.1, 1, 10, 100) are trained.

- Among the trained SVM, polynomial cubic kernel with c=1 had the best overall classification accuracy.

- The final result is the average accuracy over all k folds. The average overall classification accuracy of the proposed classification model is 96%
- F_1-score=0.96

		Predicted	
		Operational	Outage
Actual	Operational	**96.7%**	3.3%
	Outage	4.7%	**95.3%**

Conclusion

- A machine learning based outage prediction model was proposed to predict the outage of power system components based on historical event data and specific event characteristics.

- Proposed models can effectively predict outages while offering a great generalization capacity for new samples in the test subset.

- Proposed models are applicable to a variety of extreme events, and also able to consider a wide range of other features in addition to speed and component distance.

Thank you

amin.khodaei@du.edu