

# Previsão de falhas de redes inteligentes descentralizadas por aprendizado de máquina

Gabriel Arruda Evangelista\*

\*Universidade Federal do Rio de Janeiro - COPPE/PEE, Rio de Janeiro, RJ, Brasil

gabrielevangelista7@poli.ufrj.br

CPE775 - Aprendizado de Máquina

**Resumo**—Visando a estabilização de redes elétricas com a inserção de fontes renováveis foi desenvolvido o Decentral Smart Grid Control (DSGC). Esse sistema não necessita mudanças significativas na infraestrutura da rede e responde à demanda dos consumidores associando características de mercado à aspectos físicos locais da rede, de forma que há auto regulação. A literatura já apresenta modelos matemáticos do DSGC, no entanto, estas dependem de diversas hipóteses simplificadoras fortes. Uma abordagem possível e baseada em dados busca a estimativa da estabilidade das redes em função das condições de operação. A partir de dados simulados, outros trabalhos já conseguiram prever estas condições com emprego de modelos sofisticados incluindo diversas arquiteturas de redes neurais. No entanto, é possível extrair tais previsões com modelos mais simples e técnicas de aprendizado de máquina com iguais das métricas de desempenho.

**Index Terms**—Aprendizado de Máquina, Aprendizado não supervisionado, Smart Grid

## I. INTRODUÇÃO

As fontes renováveis de energia muitas vezes apresentam variações temporais e em escala. Por exemplo, turbinas eólicas podem apresentar tanto flutuações sazonais quanto em função de ventos turbulentos [1]. Portanto, é necessário que as redes sejam resilientes a esta volatilidade para atender a atual tendência de inclusão destas fontes nas matrizes energéticas.

A proposta da regulação das redes pela demanda para redução dos impactos de oscilações dos nós produtores gera uma grande mudança de paradigma do sistema de fornecimento elétrico mas vem se mostrando como uma alternativa promissora [2] [3]. A maioria das propostas nesse sentido indicavam modelos centralizados [4] para negociação de preços e adequação do fornecimento.

Se sabe que a frequência pode fornecer as informações necessárias para controle da rede. A frequência aumenta com aumento o superávit de produção de energia [5]. Fato este já utilizado no controle primário de produtores de energia atualmente. Esse fato vem sendo base para uma proposta descentralizada que pode representar menores impactos em infraestrutura e em termos segurança. Nessa contexto os consumidores podem assumir o papel de reguladores de sua demanda frente a rede elétrica.

Para isso se tornar economicamente favorável, o preço da energia observada por cada consumidor deve ser função da frequência local da rede. Existem trabalhos que modelaram sistemas descentralizados considerando preços locais

atribuindo ao sistema de controle a função de ter uma medida da estabilidade da rede [6] [7].

Um conjunto de dados, inicialmente disponibilizados pela Universidade da Califórnia, Irvine (UCI), apresenta diferentes condições de operação uma pequena rede simulada com simplificações. Uma versão aumentada [10] desta base é disponibilizada na plataforma Kaggle [8]. A partir dessa, o presente trabalho visa extrair modelos de previsão da condição de estabilidade da rede. São empregados modelos de aprendizado de máquina aplicados diretamente e posteriormente é feito um estudo com emprego de técnicas aprendizado não supervisionado para tratamento prévio dos dados.

A próxima seção descreve sucintamente o sistema estudado. Na seção seguinte a base de dados utilizada para estimativa das condições de estabilidade é apresentada. A seção IV expõe a metodologia para análise do desempenhos dos algoritmos. Na seção V os resultados são identificados e a conclusão é elaborada na seção VI.

## II. DECENTRAL SMART GRID CONTROL

Em redes inteligentes há negociação constante do preço. O consumidor indaga o preço, a rede reúne informações de consumo e fornecimento, indica o preço da energia, e, com esta resposta, o nó consumidor decide pelo uso. Como todo o processo é dependente do tempo, estimar dinamicamente a estabilidade da rede elétrica e tomar ações corretivas para mantê-la estável torna-se requisito crítico.

O objetivo das simulações é compreender os distúrbios e flutuações de produção e do consumo introduzidos pelos participantes de forma dinâmica tanto do ponto de vista técnico quanto de mercado. O principal canal de comunicação é a frequência, que indica todos estes aspectos para os participantes. Em condições de superprodução a frequência aumenta e em momentos de subprodução diminui [7]. A leitura da frequência em cada nó da rede é informação suficiente para que o administrador determine sua estabilidade e precifique a energia. Assim é possível estimar suas condições técnicas, tempo de reação ( $\tau$ ) e potência (P), e de mercado, a elasticidade preço. A elasticidade preço ( $\gamma$ , g) é composta por uma componente natural da rede e de uma uma componente de controlada pela qual a administradora intervém na rede. A componente de controle foi mantida constante nas simulações que são analisadas [6]. Portanto os dados da base representam apenas o comportamento natural da rede, sem atuação.

### III. BASE DE DADOS

O modelo matemático do DSGC é descrito em [7] e consiste de equações diferenciais de diversos aspectos físicos e do modelo de mercado proposto. Em [6] o sistema é simulado para uma rede simples como a apresentada na Figura 2 e então são realizadas perturbações. Os dados presentes na base de dados corresponde as repostas do sistema simulado e sua condição de estabilidade.

A base de dados representa essa simulação de uma rede de 4 nós, sendo um nó produtor e os demais, consumidores. Cada nó  $i$  apresenta 3 atributos, tempo de reação  $\tau_i$  (referidas com "tau1" a "tau4"), elasticidade-preço  $g_i$  (referidas como "g1" a "g4") e potência consumida ou fornecida  $P_i$  ("p1" a "p4"), totalizando 12 atributos por evento da base de dados. O nó produtor é o de índice 1 e os consumidores, de 2 a 4. Os valores de potência são positivo quando há geração de energia pelo nó ou negativo quando há consumo. Pela condição de equilíbrio vale a relação da equação 1. Os possíveis valores para cada parâmetro é dado pela Tabela I.

$$p_1 = -(p_2 + p_3 + p_4) \quad (1)$$

As variáveis  $\tau_i$  e  $g_i$  e "p2" a "p4" são distribuídas uniformemente no seu domínio enquanto a variável "p1", dependente das demais, e por isso, apresenta-se em uma distribuição normal. As distribuições das características do nó consumidor 2 e a potência do nó 1 são apresentadas na Figura 1. Os demais atributos seguem o mesmo padrão.

Tabela I  
FAIXAS DE VALORES POSSÍVEIS

Variáveis	Faixas de valores		
	Tipo de Variável	Mínimo	Máximo
tau1 a tau4	Real	0,5	10,0
g1 a g4	Real	0,05	1,00
p2 a p4	Real	-2,0	-0,5

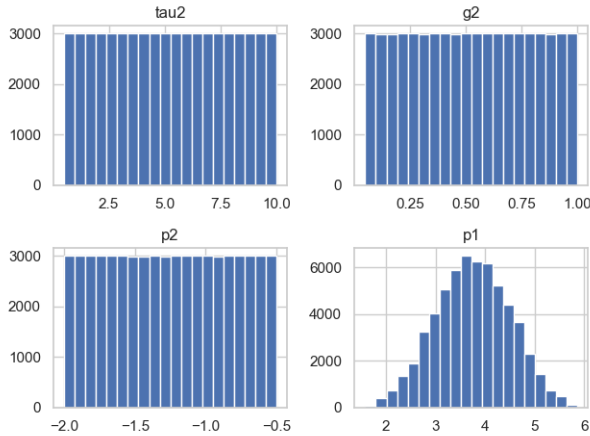


Figura 1. Distribuição dos dados.

Cada instância é classificada como estável ou não estável identificado pelo rótulo "stabf", um booleano de valor verdadeiro para condições estáveis. A base de dados apresenta 60000 instâncias sendo 63,8% rotuladas como instáveis e 36,2% como estáveis da rede, tal desbalanceamento de classes assinala a necessidade da estratificação nos processos de reamostragem. Por serem originados de uma simulação não há dados ausentes, sem haver necessidade de tratamentos para este aspecto.

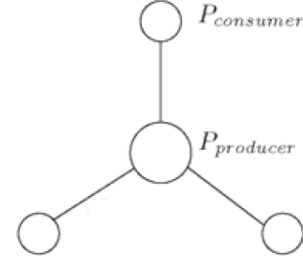


Figura 2. Modelo da rede na qual os dados foram levantados [6].

A Figura 3 apresenta a análise da matriz de correlação entre as variáveis da base de dados em um mapa de cores. A correlação apresentada é de Pearson. A correlação entre as características "p1" e "p2" a "p4" se observa no gráfico e se dá pela sua conexão devido a equação 1. Excetuada a diagonal principal e a relação entre as potência, os valores absolutos não são superiores a 0,3. Este nível de correlação faz com que técnicas de seleção de *features* impliquem perda de informação relevante.

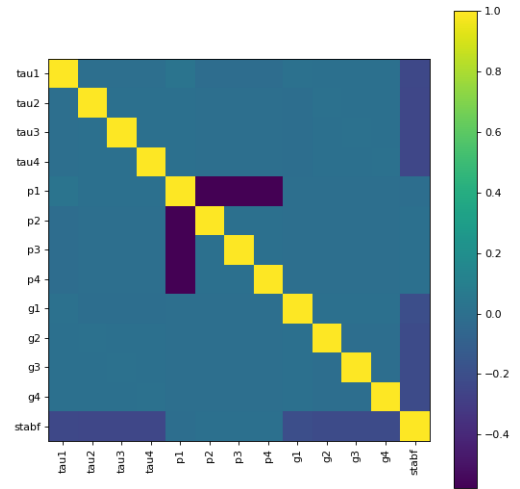


Figura 3. Correlação entre as variáveis.

### A. Trabalhos relacionados

A classificação destes dados já foi abordada antes pelos próprios autores do modelo matemático com emprego de árvores de decisão [6]. Este algoritmo foi empregado buscando manter a explicabilidade do modelo resultante. Foi relatada uma taxa de acurácia de cerca de 80%, no entanto, os parâmetros empregados não são claramente descritos e o foco inicialmente era a explicabilidade. Posteriormente [9] emprega modelo baseado em LSTMs para endereçar o problema similar, com taxas de acurácia de cerca de 99,07% e métrica F1 de 99,00%. [10] apresenta resultados de acurácia de 99,64%, com uma rede neural *feedforward* porém com 5 camadas (número de nós de cada camada: 288-288-24-12-1). Nenhum destes trabalhos apresenta as dispersões dos resultados encontrados.

## IV. METODOLOGIA

Os algoritmos testados foram aplicados empregando um K-Fold aninhado com reamostragem estratificada, sendo 10 dobras internas e 10 dobras externas. Para cada conjunto de dados foram analisadas a acurácia e a medida F1 tanto para atribuição de classificação estável quanto instável. A acurácia foi a métrica principal utilizada em [6], [9] e [10]. A medida F1 apresenta uma alternativa como indicador global associando precisão e *recall*.

Este processo de reamostragem gerou 100 amostras de validação e treino e forneceu 10 estimativas de teste de cada métrica, a partir das quais foram analisadas seus valores médios e desvios padrão. Também são apresentadas as suas distribuições em gráficos.

### A. Classificação com análise não supervisionada prévia

Em seguida foi feita uma análise com os dados tratados previamente por pela técnica fuzzy c-means [11]. Observou-se o aspecto agregado dos dados e então foi realizada a predição em função do pertencimento das amostras a cada *cluster*. Foram analisadas duas composições de dados advindos deste método conforme descrito na seção seguinte.

Neste caso, apenas dois algoritmos foram analisados a partir do mesmo processo de reamostragem, porém agora com 5 dobras externas e 5 internas.

## V. RESULTADOS

Foram empregados os seguintes algoritmos na análise dos dados:

- *K-Nearest Neighbors*
- *Adaptive Boost* (AdaBoost)
- Árvores de Decisão (CART)
- *Random Forest*
- *Gradient Boost* (LGBM, [12])

Cada um dos algoritmos foi testado mediante a metodologia exposta na seção anterior. Os parâmetros para cada algoritmo são selecionados por uma *GridSearch* entre os elencados na Tabela II nas etapas de validação do *Nested K-Fold*.

Os parâmetros relativos ao *AdaBoost* são inferiores aos empregados no *LightGBM*, o que poderia prejudicar seus indicadores de desempenho. No entanto, para parâmetros

similares, foi obtida uma acurácia média de 87,77%, não significativamente superior ao observado. O tempo de realização do teste proposto de apenas a métrica de acurácia foi de cerca de 12 horas, em comparação com cerca de 40 minutos para o *LightGBM*, sendo o processamento realizado em computador com processador i7-11800H e 16GB de memória RAM. Devido a isso, foram utilizados parâmetros menos agressivos para o *AdaBoost*.

Tabela II  
PARÂMETROS TESTADOS PARA CADA ALGORITMO

Algoritmo	Parâmetros	Valores
Árvores de Decisão	max_depth	{50, 100, 1000}
	min_samples_leaf	{1, 4, 64}
Random Forest	max_depth	{25, 50, 100}
	n_estimators	{5, 25, 100, 200}
	min_samples_leaf	{1, 4, 64}
KNN	n_neighbors	{1, 6, 9, 18}
LGBM	max_depth	{25, 50, 100}
	n_estimators	{400, 1600, 3200}
	min_child_weight	{4, 16, 64}
AdaBoost	n_estimators	{25, 50, 100, 500}
	learning_rate	{.1, 1.0, 2.0}

As figuras 4, 5 e 6 apresentam as distribuições dos 10 resultados da Medida F1 para classificação estável, da Medida F1 para classificação instável e da acurácia de cada um dos modelos no processo de reamostragem. A tabela III consolida os resultados e permite a comparação com os principais resultados de [9] e [10].

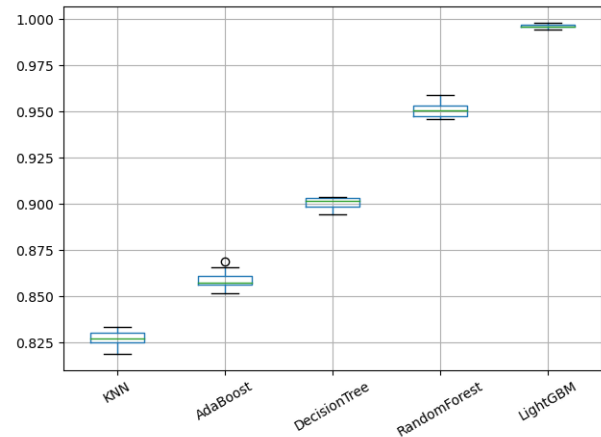


Figura 4. Distribuição das medidas F1 para classificação de condições estáveis.

O algoritmo *Gradient Boost* na implementação *Light Gradient Boosting Machine* (LGBM) [12] foi o que apresentou melhor resultado em todas as figuras de mérito, com uma taxa de acurácia de 99,61% e desvio padrão de 0,10%. Essa métrica é próxima a apresentada pela rede neural de cerca de 95000 pesos proposta por [10]. Os parâmetros que apresentaram melhor acurácia para cada algoritmo são apresentados na tabela IV.

A árvore de decisão apresentou melhor resultado do que o indicado no artigo original [6], uma possível explicação

Tabela III  
RESULTADOS CONSOLIDADOS

Algoritmo	Medida F1 estável		Medida F1 instável		Acurácia	
	Média [%]	Desvio [%]	Média [%]	Desvio [%]	Média [%]	Desvio [%]
KNN	82,71	0,42	82,75	0,53	83,05	0,37
AdaBoost	85,87	0,55	86,38	0,69	86,38	0,29
Árvore de Decisão	90,04	0,36	90,10	0,29	90,12	0,48
Random Forest	95,08	0,41	95,24	0,27	95,15	0,25
LGBM	99,61	0,10	99,62	0,09	99,62	0,09
MLSTM [9]	99,00	-	99,00	-	99,07	-
NN (288-288-24-12-1) [10]	99,69	-	-	-	99,62	-

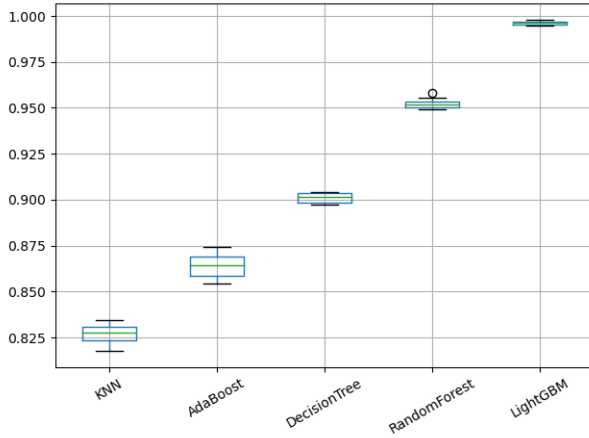


Figura 5. Distribuição das medidas F1 para classificação de condições instáveis.

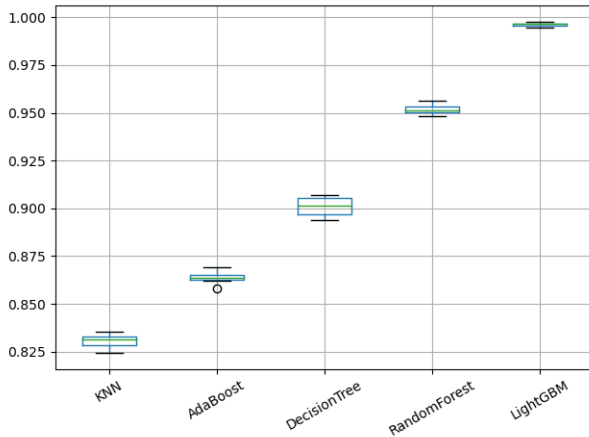


Figura 6. Distribuição das acurácias.

advém dos parâmetros empregados para algoritmo. Aqui foram utilizados valores que permitem modelos complexos. No artigo citado não são mencionados os parâmetros usados para extração do modelo da árvore mas destaca-se a importância do aspecto de explicabilidade para os autores na realização de seu trabalho, o que demanda parâmetros mais restritos.

Tabela IV  
MELHORES PARÂMETROS

Algoritmo	Parâmetros	Valores
Árvores de Decisão	max_depth	50
	min_samples_leaf	1
Random Forest	max_depth	100
	n_estimators	200
	min_samples_leaf	1
KNN	n_neighbors	9
LGBM	max_depth	25
	n_estimators	3200
	min_child_weight	4
AdaBoost	n_estimators	500
	learning_rate	1.0

#### A. Classificação com análise não supervisionada prévia

Uma possível abordagem para análise prévia dos dados é o uso da técnica de *clustering* como uma forma de extração de features. Neste sentido foi realizado experimento no qual a base de dados foi submetida ao algoritmo de clusterização suave fuzzy c-means. Os dados foram separados por nós e submetidos ao algoritmo. Foi utilizada a implementação da biblioteca skfuzzy 0.2 [11], [13] e [14]

Foi realizado um estudo exploratório para definir a quantidade adequada do número de agrupamentos, a métrica da qualidade observada foi o *fuzzy partition coefficient* (FPC). Os dados foram agrupados pelo algoritmo em 2 a 10 conjuntos e foi observado o FCP. O resultado é apresentado na Figura 7 e indica que o único número de agrupamentos que maximiza a condição de particionamento para todos os nós é 3. Destaca-se que os agrupamentos dos nós consumidores são concêntricos, o que poderia viabilizar uma análise da condição de estabilidade nó a nó e viabilizar uma implementação distribuída no sistema estudado.

Dado esse número de agrupamentos, passa-se a observar os dados após a clusterização. Os pertencimentos de cada nó a cada um dos agrupamentos passa a ser percebido como um novo atributo do registro. Dessa forma, os novos atributos nomeados ' $n_{ic_j}$ ', onde  $i$  é o nó correspondente da linha original e  $j$  o agrupamento.

Com estas novas *features* é possível fazer a matriz de correlação da função de pertencimento de cada uma das ocorrências da base de dados, o resultado é apresentado na Figura 8. Observa-se que a correlação de cada dado isoladamente é baixa, no entanto, cada conjunto relativo um nó

sempre apresentam um agrupamento de maior correlação com a variável 'stabf', que corresponde a estabilidade do sistema.

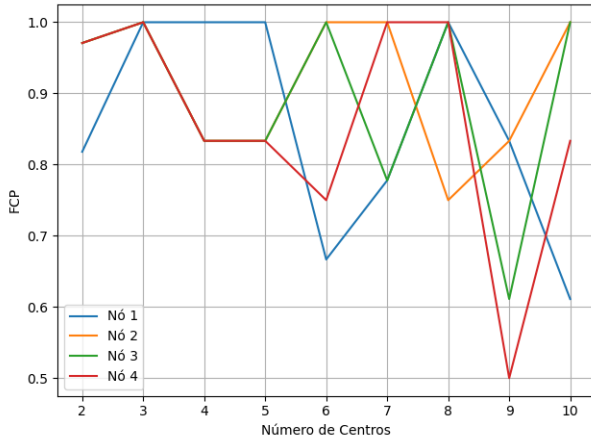


Figura 7. Figura de mérito da clusterização de cada nó em função do número de agrupamentos.

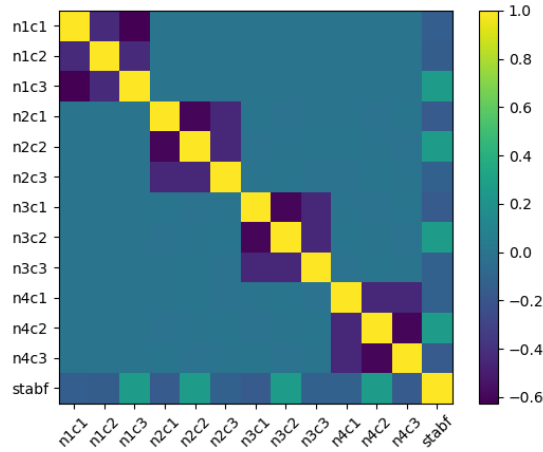


Figura 8. Correlação entre as funções de pertinência do clustering e a variável que corresponde a estabilidade da rede.

Estas novas *features* foram então analisadas de duas maneiras. Primeiramente, as pertinências foram consideradas como a única fonte de dados para predição da condição de estabilidade da rede. Em seguida elas foram empregadas justapostas aos dados iniciais. Estas duas configurações de dados foram submetidas aos testes com os algoritmos de Árvores de Decisão e LightGBM. Os resultados foram apresentados na Tabela V. Com os dados das pertinências somente ("Pert." na tabela) ainda é possível realizar o processo de classificação, no entanto, os resultados apresentados foi pior do que o emprego direto do algoritmos. Observa-se, que quando justapostos ("Just." na tabela) o impacto nas métricas de desempenho é uma ligeira perda, possivelmente irrelevante, podendo ser considerado indiferente.

Tabela V  
RESULTADOS APÓS CLUSTERIZAÇÃO

		Acurácia		F1 (Estável)	
		Média [%]	Desvio [%]	Média [%]	Desvio [%]
Pert.	Árvore D.	76,75	0,92	76,36	0,22
	LGBM	87,52	0,19	87,52	0,20
Just.	Árvore D.	89,24	0,24	89,23	0,20
	LGBM	99,52	0,07	99,52	0,07

## VI. CONCLUSÃO

Este trabalho analisou o desempenho de diversos estimadores para a abordagem do problema de identificação da estabilidade de redes DSGC e abordou um processo alternativo tratando inicialmente os dados com técnica não supervisionada.

A partir dos resultados apresentados se depreende que os algoritmos de *gradient boosting*, particularmente a implementação LightGBM, tem capacidade de apresentar mesmo nível de precisão que redes neurais apresentadas em outros trabalhos para o problema analisado mediante a configurações de parâmetros suficientemente robustos.

A análise dos clusters permitiu observar que todos os nós se agrupam de igual maneira em 3 *clusters*. Assim, o tratamento com o método proposto poderia permitir uma análise das redes nó a nó, com potencial para viabilizar ganhos de escala ou implementações distribuídas para tratar redes maiores. Porém, no problema tratado, apesar de viabilizar a identificação da condição de estabilidade, a técnica não apresentou ganhos nas métricas analisadas. A análise apenas dos dados das pertinências implicou métricas piores e também não houve ganho precisão quando empregada aglutinadas com dados original.

## REFERÊNCIAS

- [1] D. Heide, L. Von Bremen, M. Greiner, C. Hoffmann, M. Speckmann, and S. Bofinger, "Seasonal optimal mix of wind and solar power in a future, highly renewable europe," *Renewable Energy*, vol. 35, no. 11, pp. 2483–2489, 2010.
- [2] D. Butler, "Energy efficiency: Super savers: Meters to manage the future," *Nature*, vol. 445, no. 7128, pp. 586–589, 2007.
- [3] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE transactions on industrial informatics*, vol. 7, no. 3, pp. 381–388, 2011.
- [4] J. K. Kok, C. J. Warmer, and I. Kamphuis, "Powermatcher: multiagent control in the electricity infrastructure," in *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, 2005, pp. 75–82.
- [5] J. A. Short, D. G. Infield, and L. L. Freris, "Stabilization of grid frequency through dynamic demand control," *IEEE Transactions on power systems*, vol. 22, no. 3, pp. 1284–1293, 2007.
- [6] V. Arzamasov, K. Böhm, and P. Jochem, "Towards concise models of grid stability," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2018, pp. 1–6.
- [7] B. Schäfer, C. Grabow, S. Auer, J. Kurths, D. Witthaut, and M. Timme, "Taming instabilities in power grid networks by decentralized control," *The European Physical Journal Special Topics*, vol. 225, no. 3, pp. 569–582, 2016.
- [8] P. Breviglieri, "Smart grid stability, <https://www.kaggle.com/datasets/pcbreviglieri/smart-grid-stability>," online, acessado em 18/09/2022.

- [9] M. Alazab, S. Khan, S. S. R. Krishnan, Q.-V. Pham, M. P. K. Reddy, and T. R. Gadekallu, "A multidirectional lstm model for predicting the stability of a smart grid," *IEEE Access*, vol. 8, pp. 85 454–85 463, 2020.
- [10] P. Breviglieri, T. Erdem, and S. Eken, "Predicting smart grid stability with optimized deep models," *SN Computer Science*, vol. 2, no. 2, pp. 1–12, 2021.
- [11] R. Winkler, F. Klawonn, and R. Kruse, "Fuzzy c-means in high dimensional spaces," *International Journal of Fuzzy System Applications (IJFSA)*, vol. 1, no. 1, pp. 1–16, 2011.
- [12] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, vol. 30, 2017.
- [13] T. J. Ross, *Fuzzy logic with engineering applications*. John Wiley & Sons, 2005.
- [14] J. C. Bezdek, C. Coray, R. Gunderson, and J. Watson, "Detection and characterization of cluster substructure i. linear structure: Fuzzy c-lines," *SIAM Journal on Applied Mathematics*, vol. 40, no. 2, pp. 339–357, 1981.