

Inferência Bayesiana Aplicada a MLPs para Previsão Probabilística de Carga Semanal

Douglas A. M. Barbosa*. Vitor H. Ferreira.**

*Universidade Federal Fluminense, Niterói, Rio de Janeiro
(Tel: 21-96467-2284; e-mail: douglasaranil@id.uff.br)

** Universidade Federal Fluminense, Niterói, Rio de Janeiro
(Tel: 21-99969-8728; e-mail: vhferreira@id.uff.br).

Abstract: An essential factor for the safe and optimized planning and operation of an electrical power system is a previous knowledge of the load during a certain period, ranging from a few minutes to years ahead. Thus, over the last years, this theme has become the target of a lot of research, and different methods were proposed. Generally, the methods created are divided into two main fields: artificial intelligence methods and statistical methods. In recent years, artificial intelligence methods have been highlighted due to their capacity to deal with nonlinear patterns, as happens with the load and their exogenous variables. Thus, this article presents the application of Bayesian Inference based neural network as load forecasting model (one week ahead).

Resumo: Um fator importantíssimo para o planejamento e operação segura e otimizado de um sistema elétrico de potência é o conhecimento a priori do valor de carga que irá se realizar num determinado horizonte de tempo, sendo esse horizonte variável desde alguns minutos até anos à frente. Com isso, ao longo dos últimos anos esse tema se tornou alvo de inúmeras pesquisas e diversos métodos de apoio a tomada de decisão foram desenvolvidos. De maneira geral, os métodos desenvolvidos se dividem em duas áreas, que são: métodos de inteligência artificial e métodos estatísticos. Ao longo dos últimos anos os métodos de inteligência artificial vêm se destacando, devido a sua grande capacidade de lidar com padrões não lineares, conforme acontece com a carga e suas variáveis exógenas. Dessa forma, esse artigo apresenta a utilização de Inferência Bayesiana para especificação de redes neurais como forma de prever a carga de energia semanal.

Keywords: Neural Networks, Machine Learning, Load Forecasting, Bayesian neural networks, Transmission System Operators.

Palavras-chaves: Redes Neurais, Aprendizado de Máquinas, Previsão de Carga, Redes Neurais Bayesianas, Operação de sistema elétricos.

1. INTRODUÇÃO

Atualmente, a previsão de carga ao longo de um determinado período vem se tornando cada vez mais importante para uma operação segura e otimizada dos sistemas elétricos de potência. Diversos estudos são embasados pelo valor previsto da carga, como por exemplo, despacho econômico, programação da geração e programação de interrupções para manutenção. Além disso, com a privatização do setor elétrico em diversos países, a energia elétrica se tornou uma commodity, sendo comercializada em mercados de energia (Hippert, et al., 2001). Como a carga representa uma das variáveis mais importantes para a composição do valor da energia, o seu conhecimento prévio e preciso representa um fator importantíssimo para diversos agentes do setor elétrico.

De modo geral, as previsões de carga apresentam três tipos de horizontes a serem considerados: curto prazo (STFL), médio prazo (MLTF) e longo prazo (LTFL). Contudo, as definições desses horizontes não são únicas, com isso previsões consideradas de longo prazo para algumas empresas podem ser

consideradas de médio para outras (Tao Hong, 2016) e (Ferreira, 2008).

Embora, diversos modelos de previsão tenham sido apresentados na literatura nos últimos anos, essa atividade ainda é um grande desafio para seus idealizadores, podendo trazer prejuízos de milhões de reais (Douglas, et al., 1998). De forma geral, os modelos apresentados na literatura podem ser divididos em duas áreas: modelos estatísticos de análise de séries temporais (ARMA, ARIMA, SARIMA) e os modelos de inteligência artificial (Redes Neurais, Logica Fuzzy, Deep Learning, Neuro Fuzzy) (Shyh-Jier Huang and Kuang-Rong Shih, 2003), (de Andrade and da Silva, 2009), (Hanmandlu and Chauhan, 2011), (Kwang-Ho Kim et al., 1995), (Amin and Hoque, 2019) e (Fan and Chen, 2006). Na maioria dos casos, os modelos estatísticos de séries temporais apresentam bons resultados quando associados aos problemas de previsão de carga, todavia resultados espúrios são observados quando existem mudanças de cenários como por exemplo: a variação brusca da temperatura entre duas horas (Zor, et al., 2017).

A limitação dos métodos de séries temporais adicionada ao avanço do poder computacional são fatores que corroboram para o crescimento do número de aplicações de modelos que utilizam inteligência artificial (IA). Dentre os principais modelos de IA aplicados para previsão de carga a curto e médio prazo, as redes neurais artificiais vêm apresentando destaque (Ferreira & Alves da Silva, 2007). Contudo, a maioria das aplicações ainda requer a intervenção de especialistas para especificação de alguns hiperparâmetros como, por exemplo, o número de neurônios na camada escondida.

Para superar estas limitações, este artigo apresenta um modelo para previsão probabilística de carga semanal utilizando Inferência Bayesiana para especificação e treinamento de perceptrons de múltiplas camadas, popularmente conhecidos como MLPs (do inglês *Multi-layer Perceptron*). Dentre as principais vantagens da metodologia podem ser citados:

- Seleção automática das variáveis relevantes de entrada;
- Utilização de todo conjunto de dados para o treinamento do modelo devido ao controle da complexidade realizado pelo algoritmo de treinamento.

Para o desenvolvimento do modelo foram utilizados dados de carga semanal de todos os subsistemas que compõem o Sistema Interligado Nacional (SIN). Essas informações foram disponibilizadas pelo Operador Nacional do Sistema Elétrico (ONS).

2. PROGRAMA MENSAL DA OPERAÇÃO

No Brasil, o ONS é o órgão responsável pela coordenação e controle da operação das instalações de transmissão e geração do país. Além disso, também está sob sua responsabilidade o planejamento da operação dos sistemas isolados, sob a fiscalização e regulação da Agência Nacional de Energia Elétrica (ANEEL) (Operador Nacional do Sistema Elétrico, 2019). A Figura 1 apresenta os subsistemas que compõem o SIN.

De forma a cumprir suas obrigações legais e missão institucional, o ONS desenvolve uma série de estudos e ações para garantir a operação segura, otimizada e contínua em todo país. Uma dessas ações desenvolvidas é o Programa Mensal da Operação (PMO), o qual visa elaborar as diretrizes eletroenergéticas para uma operação otimizada, segura e com o menor custo possível, atendendo os padrões técnicos e os critérios de confiabilidade estabelecidos nos Procedimentos de Rede.

Um dos insumos do PMO é a previsão de carga de curto prazo. No submódulo 5.6 dos procedimentos de rede são definidas as responsabilidades dos agentes e do próprio ONS para que ocorra a consolidação da previsão de energia média (MW_{médio}) e demanda máxima (MW) para os subsistemas e o SIN. Nesse submódulo encontra-se definido que devem ser enviadas ao ONS as previsões de carga em discretização mensal para o horizonte no qual o PMO é realizado. Além das

previsões mensais, para o primeiro mês de previsão a carga prevista deverá ser desagregada em semanas operativas do mês, como apresentado na Figura 2 (Operador Nacional do Sistema Elétrico, 2018). Os agentes utilizam diferentes metodologias para elaboração de suas previsões. Neste trabalho serão utilizadas redes neurais artificiais conforme apresentado na próxima seção.



Figura 1 - Subsistemas que compõem o Sistema Interligado Nacional

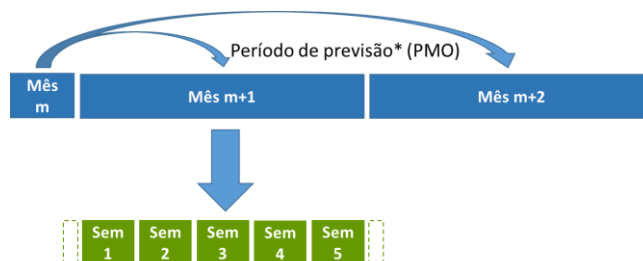


Figura 2 - Horizontes de Previsão do PMO

3. REDES NEURAIS ARTIFICIAIS

Atualmente, um dos principais modelos de IA são as redes neurais artificiais (RNAs), as quais são modelos matemáticos que possuem como objetivo simular o funcionamento do cérebro humano (Martins, 2018). Suas principais vantagens de utilização são:

- Não linearidade: Usualmente, as funções de ativação dos neurônios da camada oculta são não lineares, com isso o modelo resulta num considerável grau de não linearidade entre o conjunto de entrada e sua saída.
- Mapeamento Entrada-Saída: Dado um conjunto de dados que relaciona entrada com saídas, as RNA conseguem relacionar tais informações sem a necessidade de definição de uma função específica.
- Adaptabilidade: As RNAs possuem elevada capacidade de se adaptar às informações que são apresentadas, para isto basta treinar, novamente, o

modelo para que haja um reajuste a partir das novas informações inseridas.

- **Simplicidade:** Com a interpretação dos modelos *feedforward* como grafos orientados, a implementação dessas estruturas é extremamente simples quando comparada ao grau do modelo que pode ser gerado.

Embora sua implementação seja extremamente simples, existem diversos hiperparâmetros que necessitam serem ajustados durante sua utilização. Este fato pode dificultar a aplicação desta técnica a problemas envolvendo a previsão simultânea de múltiplas séries temporais, inviabilizando a especificação dedicada de cada modelo. Normalmente, os dois principais problemas relacionados com a especificação do modelo são: escolha das variáveis de entrada e controle de complexidade do modelo, visando obter a precisão requerida e evitando o *overfitting* (Ferreira & Alves da Silva, 2007).

As RNAs são capazes de aproximar com precisão arbitrária qualquer função contínua (Zang, et al., 1998), desde que seja escolhido o número ideal de neurônios na camada escondida. A definição deste hiperparâmetro é uma das etapas do controle de complexidade. Na literatura são apresentados diversos métodos para esse controle, tais como:

- **Pruning Algorithms:** o modelo começa com um determinado número de neurônios, com esse parâmetro sendo reduzido de forma iterativa. Ao final é escolhido o modelo mais simples e com menor erro, obedecendo assim o princípio da parcimônia (Hippert & Taylor, 2010).
- **Regularization techniques:** Esta técnica adiciona um termo de penalização na função objetivo de treinamento, logo modelos com valores de pesos muito elevados são penalizados e não escolhidos ao final (Hippert & Taylor, 2010).

Além do problema de escolha do número ideal de neurônios na camada escondida, existe também o desafio de se obter o melhor conjunto de entradas a serem utilizadas. Diversos trabalhos, não definem uma metodologia para escolha dessas variáveis, somente informando quais dessas foram utilizadas. Isso resulta em uma perda de capacidade de generalização do modelo, uma vez que sempre será necessário a visão de um especialista para julgar qual é a melhor variável a ser utilizada no problema, sendo isso algo indesejado.

Com a escolha das variáveis de entrada e do número de neurônios a serem utilizados, existe a necessidade do ajuste de diversos parâmetros da rede, como seus pesos e bias. Muitos autores utilizam a técnica de validação cruzada (VC) para a verificação do ajuste do modelo. Porém, essa técnica apresenta a necessidade da separação do conjunto de dados disponíveis em dois, uma para o treinamento e outro para validação. Logo, inúmeros modelos precisam ser treinados para que se escolha aquele que apresentar o menor erro para o conjunto de validação. Contudo, muitos autores ainda se apresentam bastante céticos durante a utilização dessa técnica, uma vez que nada garante que o modelo escolhido é o melhor deles (Catatelpo, et al., 1999). A técnica de VC ainda apresenta outra limitação, pois caso seja utilizado diferentes formas de

validação (*leave-one-out*, *K-fold*, *bootstrap*) cada uma delas podem apresentar resultados distintos durante a escolha do modelo ideal. (Efron & Tibshirani, 2016).

Para tratar estas questões (seleção de entradas e controle de complexidade) de forma automática e acoplada, este trabalho apresenta a aplicação de Inferência Bayesiana ao treinamento e especificação de redes neurais do tipo MLP. Esta técnica foi proposta por *David J.C. Mackay* (Mackay, 1992) dando origem a um algoritmo de treinamento de MLPs que inclui mecanismos automáticos para seleção de entradas e controle de complexidade conforme será apresentado na próxima seção.

3.1 Inferência Bayesiana Aplicada aos MLPs

Ao contrário dos demais treinamentos aplicados às RNAs, os quais possuem, em geral, um único objetivo de minimização do erro quadrático, o treinamento bayesiano possui como objetivo maximizar a probabilidade a posteriori do conjunto de pesos dado os padrões de entrada e saída. Para isso, utiliza-se como auxílio a regra de Bayes para determinar a probabilidade a posteriori $p(\underline{W} | X, D)$, dada por:

$$p(\underline{W} | X, D) = p(D | X, \underline{W}) \frac{p(\underline{W})}{p(D, X)} \quad (1)$$

onde $p(D | X, \underline{W})$ representa a função de verossimilhança, $p(\underline{W})$ a probabilidade a priori de ocorrência do vetor de pesos \underline{W} e $p(D, X)$ a probabilidade de ocorrência de D dado X .

Como a $p(D, X)$ não depende de \underline{W} , esta probabilidade atua como um fator de normalização que não irá impactar no problema de maximização de $p(\underline{W} | X, D)$. Portanto, o desafio é especificar $p(D | X, \underline{W})$ e $p(\underline{W})$, dadas por (Ferreira & Alves da Silva, 2007):

$$p(\underline{W}) = \frac{1}{\sqrt{\frac{2\pi}{\alpha}}^M} e^{-\frac{\alpha}{2} \sum_{i=1}^M w_i^2} \quad (2)$$

$$p(D | X, \underline{W}) = \frac{1}{\sqrt{\frac{2\pi}{\beta}}^N} e^{-\frac{\beta}{2} \sum_{i=1}^N \{d^{(k)} - f[x^{(k)}, \underline{W}]\}_i^2} \quad (3)$$

A probabilidade a priori $p(\underline{W})$ parte da premissa que inicialmente os pesos devem assumir valores próximos a zero visando evitar a saturação do treinamento, em analogia à normalização aplicada aos sinais de entrada e saída (Ferreira, 2008). Assim, é adotado que o conjunto de pesos segue uma distribuição normal com média nula e matriz de covariância proporcional à identidade (αI).

A verossimilhança $p(D | X, \underline{W})$ assume que os erros são independentes e distribuídos identicamente segundo uma distribuição normal com média nula e variância $1/\beta$. Estas

especificações dão origem ao problema de otimização dado por:

$$\begin{aligned} \max_{\underline{W}} p(\underline{W} | X, D) &= \max_{\underline{W}} p(D | X, \underline{W}) p(\underline{W}) = \\ &= \max_{\underline{W}} \frac{1}{\sqrt{\alpha\beta}^{M+1}} e^{-\left[\frac{\beta}{2} \sum_{i=1}^N \{d(k) - f[x(k), \underline{W}]\}_i^2 + \frac{\alpha}{2} \sum_{j=1}^M W_j^2 \right]} \end{aligned} \quad (4)$$

Para maximizar $p(D | X, \underline{W})$ é necessário é minimizar o expoente da expressão (4), ou seja:

$$\min_{\underline{W}} \left\{ \frac{\beta}{2} \sum_{i=1}^N \{d(k) - f[x(k), \underline{W}]\}_i^2 + \frac{\alpha}{2} \sum_{j=1}^M W_j^2 \right\} \quad (5)$$

A equação (5) constitui a métrica a ser minimizada ao se utilizar o treinamento Bayesiano, que inclui ainda métodos analíticos para estimativa de β e α (Mackay, 1992). Nela o primeiro termo representa a capacidade do modelo de se ajustar aos dados de treino, já a segunda parcela garante o controle de complexidade do modelo, evitando assim o problema de ajuste excessivo aos dados por meio do controle da magnitude dos pesos. Assim, pesos em excesso, erroneamente especificados por meio de entradas irrelevantes ou por um número elevado de neurônios na camada escondida, terão valores estimados próximos a zero promovendo o chamado *soft pruning*. De fato, a escolha de uma probabilidade a priori $p(\underline{W})$ independente para grupos de pesos distintos permite dividir um método para determinação automática de relevância conhecido como ARD (*automatic relevance determination*). Maiores detalhes podem ser encontrados em (Ferreira & Alves da Silva, 2007).

4. RESULTADOS

No Brasil, conforme supracitado, existem 4 subsistemas que compõem o SIN. Embora cada um desses subsistemas apresente suas peculiaridades e características, foram escolhidas entradas semelhantes para todos os subsistemas, dado que é esperado que o método ARD proposto atue caso algumas dessas variáveis sejam irrelevantes para o subsistema em questão. O número de neurônios na camada escondida foi mantido constante igual a 5, com o treinamento bayesiano sendo responsável pelo controle de complexidade. Os dados utilizados como entradas iniciais foram: carga e temperatura (máxima, média e mínima) das 4 últimas semanas, em relação à semana alvo de previsão, variável binária para representação de feriados e temperatura prevista para à semana alvo de previsão.

Utilizou-se como período para treinamento do modelo o intervalo entre a primeira semana operativa de 2014 (28/12/2013 – 03/01/2014) e a penúltima semana operativa de 2016 (24/12/2016 – 30/12/2016). Para teste do modelo foi utilizado o intervalo entre a primeira semana operativa de 2017 (31/12/2016 – 06/01/2017) e a última semana operativa de 2018 (29/12/2018 – 04/01/2019).

Como forma de avaliação do modelo foi utilizado o MAPE (*Mean Absolut Percentage Error*), dado por:

$$Mape = \frac{1}{N} \sum_{i=1}^N \frac{V_i - P_i}{V_i} 100 [\%] \quad (6)$$

Onde V_i representa o valor verificado e P_i o valor previsto.

A seguir será apresenta os resultados obtidos utilizando a metodologia proposta, para a previsão de carga de energia (MW médios) uma semana à frente.

4.1 Estudo de caso

Ao se utilizar a metodologia proposta foi obtido um erro médio para o SIN de aproximadamente 1,88%. A Tabela 1 apresenta os erros médios (MAPEs) por subsistemas para o período de teste.

Tabela 1 - Erro médio por subsistema (cenário 1)

Subsistema	Erro (%)
SE/CO	2,11
SUL	3,29
NORDESTE	2,41
NORTE	1,90

Nas figuras 3, 4, 5, 6 e 7 são apresentadas as distribuições dos erros obtidos nesse cenário. Através desses gráficos, observa-se que a maioria dos erros (85% para o SIN, vide figura 3) se encontram até 3%, principalmente no subsistema Sudeste/Centro-Oeste (83% - figura 4), que representa o principal vetor de erro para o total do SIN.

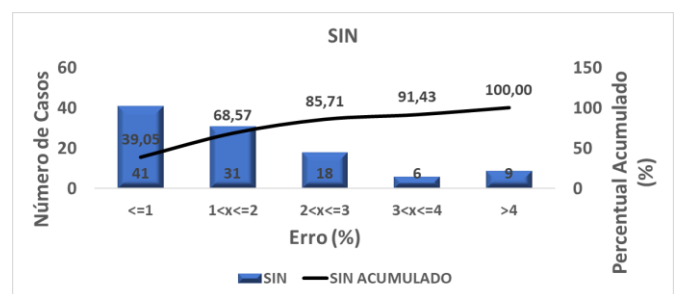


Figura 3 - Erro acumulado (SIN)

Três formas de Benchmark podem ser avaliadas para validação dos erros apresentados, a primeira seria a utilização do método Naive de previsão como comparação, onde o mesmo valor verificado de carga da semana anterior é utilizado como previsão (Hippert & Taylor, 2010). Nesse caso o erro encontrado para o SIN foi de 2,27% se mostrando assim superior ao erro verificado no presente artigo, que foi de 1,88%, justificando-se, a importância do desenvolvimento deste modelo.

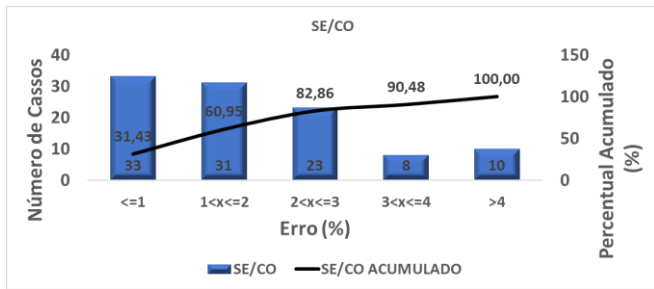


Figura 4 - Erro acumulado (Sudeste/Centro-Oeste)

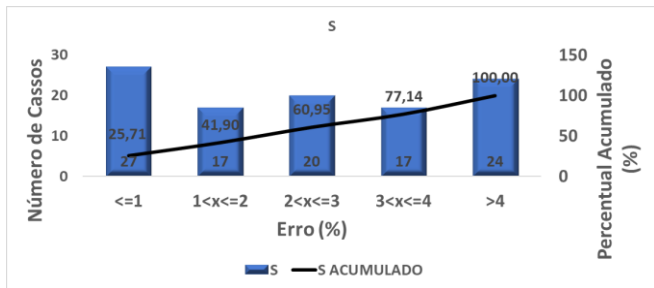


Figura 5 - Erro acumulado (Sul)

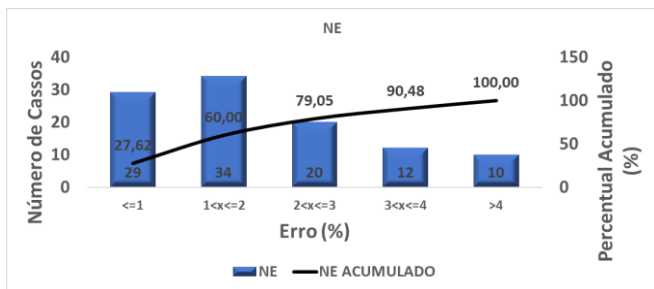


Figura 6 - Erro acumulado (Nordeste)

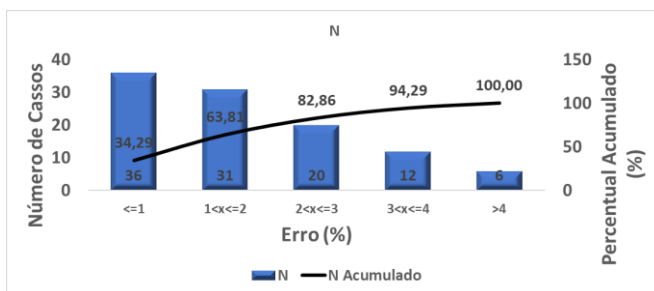


Figura 7 - Erro acumulado (Norte)

Como supracitado, modelos estatísticos apresentam um determinado atraso para ajustar sua resposta à mudança no comportamento da carga, ao contrário das redes neurais. Esse fato se torna visível na semana operativa que termina no dia 06/06/2018, onde é observado uma elevada variação no comportamento da carga, impulsionado principalmente pela variação dos valores de temperaturas. Com isso o método Naive apresentou um erro médio de aproximadamente 10% no SIN, já as redes neurais apresentaram um desvio de aproximadamente apenas 2,3%, justificando o que foi afirmado a respeito das vantagens de utilização das RNAs. A

segunda forma de Benchmark utiliza-se dos indicadores de desempenho do ONS, os quais apresentam como meta, para a gerência de previsão de carga, desvios inferiores a 3% no SIN (Operador Nacional do sistema Elétrico, 2019). Embora esse indicador diga a respeito dos desvios da carga Mensal de energia, foi adotado a mesma premissa para a carga semanal, à vista disso, conclui-se que o modelo desenvolvido estaria cumprindo com louvor a meta estabelecida, uma vez que o desvio no SIN foi inferior a 3%. Por fim, a última forma de Benchmark apresentada, consiste na comparação dos desvios semanais dos meses de Outubro/18 e Novembro/18, apresentados pelo ONS durante as reuniões do PMO.

Tabela 2 - Erros Médios Modelo Proposto (%)

MÊS	SE/CO	S	NE	N
Out	2,3	1,3	2,6	1,4
Nov	2,1	2,5	1,1	1,3

Tabela 3 - Erros Médios ONS (%) (Operador Nacional do Sistema Elétrico, 2018)

MÊS	SE/CO	S	NE	N
Out	1,0	1,0	1,7	1,5
Nov	2,3	4,4	2,0	2,1

Analisando os erros apresentados nas Tabela 2 e Tabela 3, verifica-se que o modelo proposto apresenta um desvio semelhante a equipe de previsão de carga do ONS. Fato esse que corrobora a viabilidade do desenvolvimento e utilização do modelo proposto.

5. CONCLUSÃO

Este artigo teve como objetivo apresentar a utilização da inferência bayesiana para treinamento e especificação de redes neurais do tipo MLP aplicadas ao problema de previsão da energia média semanal, uma semana à frente. Pelos resultados obtidos, onde o erro médio ficou inferior de 2% para o SIN, pôde-se verificar que o treinamento Bayesiano aplicado aos MLPs se destaca como uma forma adequada de modelagem quando inserida em problemas de previsão de carga semanal. Além disso, a metodologia trata do controle automático de complexidade do modelo e da escolha das variáveis de entradas relevantes, o que garante uma maior flexibilidade e uma menor intervenção por parte dos especialistas, conforme verificado ao apresentar resultados satisfatórios para todos os subsistemas sem requerer intervenções para ajustes tanto do conjunto de entradas, quanto do número de neurônios na camada escondida. Dado a possibilidade de uma interpretação estatística dos valores de carga de energia previstos, como trabalhos futuros pretende-se incluir intervalos de confiança associados a essas previsões e cenários de temperatura especialmente para as previsões em múltiplas semanas a frente, agregando esta informação aos intervalos de confiança para as previsões tendo por objetivo avaliar o nível de incerteza a partir das variações nas condições climáticas.

AGRADECIMENTOS

Agradeço ao Operador Nacional do Sistema Elétrico, em especial a Gerência de Previsão de Carga pela disponibilização das informações utilizadas para o desenvolvimento do presente artigo.

REFERÊNCIAS

- Amin, M.A.A., Hoque, Md.A., 2019. Comparison of ARIMA and SVM for Short-term Load Forecasting, in: 2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON). Presented at the 2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON), IEEE, Jaipur, India, pp. 1–6. <https://doi.org/10.1109/IEMECONX.2019.8877077>
- Cataltepe, Zehra, Yaser S. Abu-Mostafa, e Malik Magdon-Ismail. “No Free Lunch for Early Stopping”. *Neural Computation* 11, no 4 (maio de 1999): 995–1009. <https://doi.org/10.1162/089976699300016557>.
- de Andrade, L.C.M., da Silva, I.N., 2009. Very Short-Term Load Forecasting Based on ARIMA Model and Intelligent Systems, in: 2009 15th International Conference on Intelligent System Applications to Power Systems. Presented at the 2009 15th International Conference on Intelligent System Applications to Power Systems (ISAP), IEEE, Curitiba, pp. 1–6. <https://doi.org/10.1109/ISAP.2009.5352829>
- Douglas, A.P., A.M. Breipohl, F.N. Lee, e R. Adapa. “Risk Due to Load Forecast Uncertainty in Short Term Power System Planning”. *IEEE Transactions on Power Systems* 13, no 4 (novembro de 1998): 1493–99. <https://doi.org/10.1109/59.736296>.
- Efron, B. & Tibshirani, R. J., 2016. “An introduction to the bootstrap”. Softcover reprint of the original 1st ed. 1993 ed. Londres: CRC Press.
- Fan, S., Chen, L., 2006. Short-Term Load Forecasting Based on an Adaptive Hybrid Method. *IEEE Trans. Power Syst.* 21, 392–401. <https://doi.org/10.1109/TPWRS.2005.860944>
- Ferreira, V. H., 2008. “Desenvolvimento de Modelos Neurais Autônomos Para Previsão de Carga Elétrica”. Tese de Doutorado, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Ferreira, Vitor Hugo, e Alexandre P. Alves da Silva. “Toward Estimating Autonomous Neural Network-Based Electric Load Forecasters”. *IEEE Transactions on Power Systems* 22, no 4 (novembro de 2007): 1554–62. <https://doi.org/10.1109/TPWRS.2007.908438>.
- Hanmandlu, M., Chauhan, B.K., 2011. Load Forecasting Using Hybrid Models. *IEEE Trans. Power Syst.* 26, 20–29. <https://doi.org/10.1109/TPWRS.2010.2048585>
- Hippert, Henrique S., e James W. Taylor. “An Evaluation of Bayesian Techniques for Controlling Model Complexity and Selecting Inputs in a Neural Network for Short-Term Load Forecasting”. *Neural Networks* 23, no 3 (abril de 2010): 386–95. <https://doi.org/10.1016/j.neunet.2009.11.016>.
- Hippert, H.S., C.E. Pedreira, e R.C. Souza. “Neural Networks for Short-Term Load Forecasting: A Review and Evaluation”. *IEEE Transactions on Power Systems* 16, no 1 (Fevereiro de 2001): 44–55. <https://doi.org/10.1109/59.910780>.
- Hong, Tao, e Shu Fan. “Probabilistic Electric Load Forecasting: A Tutorial Review”. *International Journal of Forecasting* 32, no 3 (julho de 2016): 914–38. <https://doi.org/10.1016/j.ijforecast.2015.11.011>.
- Kwang-Ho Kim, Jong-Keun Park, Kab-Ju Hwang, Sung-Hak Kim, 1995. Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems. *IEEE Trans. Power Syst.* 10, 1534–1539. <https://doi.org/10.1109/59.466492>
- Mackay, D. J., 1992. “Bayesian interpolation”. *Neural Computation* 4, 415–447. <https://doi.org/10.1162/neco.1992.4.3.415>
- Martins, R. P., 2018. Previsão Probabilística de Energia Solar Utilizando Redes Neurais Bayesianas e Mistura de Gaussianas, Dissertação de Mestrado, Universidade Federal Fluminense, Niterói.
- Operador Nacional do Sistema Elétrico, 2018. “Conhecimento: Acervo Digital”. [Online] Available at: <http://www.ons.org.br/paginas/conhecimento/acervo-digital/documentos-e-publicacoes?categoria=Relat%C3%B3rio+PMO> [Acesso em 18 janeiro 2020].
- Operador Nacional do Sistema Elétrico, 2018. “Consolidação da previsão de carga para a elaboração do Programa Mensal da Operação”, Rio de Janeiro: s.n.
- Operador Nacional do sistema Elétrico, 2019. “Indicadores de Desempenho ONS”. [Online] Available at: http://www.ons.org.br/Paginas/sobre-ons/indicadores-ons/indicador_ipc.aspx [Acesso em 13 07 2019].
- Operador Nacional do Sistema Elétrico, 2019. “Sobre o ONS: O que é ONS”. [Online] Available at: <http://www.ons.org.br/paginas/sobre-o-ons/o-que-e-ons> [Acesso em 30 dezembro 2019].
- Shyh-Jier Huang, Kuang-Rong Shih, 2003. Short-term load forecasting via ARMA model identification including non-gaussian process considerations. *IEEE Trans. Power Syst.* 18, 673–679. <https://doi.org/10.1109/TPWRS.2003.811010>
- Zhang, Guoqiang, B. Eddy Patuwo, e Michael Y. Hu. “Forecasting with Artificial Neural Networks”: *International Journal of Forecasting* 14, no 1 (março de 1998): 35–62. [https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7).
- Zor, Kasim, Oguzhan Timur, e Ahmet Teke. “A State-of-the-Art Review of Artificial Intelligence Techniques for Short-Term Electric Load Forecasting”. In 2017 6th International Youth Conference on Energy (IYCE), 1–7. Budapest, Hungary: IEEE, 2017. <https://doi.org/10.1109/IYCE.2017.8003734>.