

Uso Machine Learning em Dados de EEG para Controle de Dispositivos Remotos

Jonatha Bizerra Silva

Instituto Santos Dumont

Instituto Internacional de Neuroengenharia Edmond e Lily Safra

Brasil, Rio Grande do Norte

Email: jonatha.silva@edu.isd.org.br

Abstract—The use of machine learning for prediction the disease is one of various applications of artificial intelligence. In this job, there is the implementation of two algorithms of machine learning with data of electroencephalography for predict different states of the patients. The objective in studding this implementation is use her in a dispositive that will use to control remotes dispositives.

I. PROJETO DE MESTRADO

O projeto de mestrado ainda está em desenvolvimento, mas objetiva-se a criação de um dispositivo de interface cérebro máquina para controle de equipamentos remotos. Para isso, serão utilizados dados de eletroencefalograma e um dispositivo que realizará a leitura dos eletrodos do EEG e a predição da função desejada. Para isso, serão estudadas as implementações de diferentes algoritmos para treinamento supervisionado. Quando o dispositivo estiver treinado com um nível de acurácia aceitável, o usuário poderá contralar diferentes dispositivos remotos utilizando apenas os sinais de eletroencefalograma, sem necessidade de movimentação dos membros do corpo.

II. ARTIGO ESCOLHIDO

O cérebro humano possui milhões de neurônios responsáveis pelo controle do comportamento humano, em resposta aos inúmeros estímulos sensoriais externos. A atividade elétrica desses neurônios pode ser captadas por meio de eletrodos corretamente posicionados em regiões do cérebro. Dados de eletroencefalograma foram utilizados na detecção de ataques epiléticos por Liu, XI, Zhao e Li (2019). Os pesquisadores aplicaram de algoritmos de aprendizagem de máquina tradicionais, como Regressão Logística e KNN para detectar estados de crise epiléticas. Liu, XI, Zhao e Li (2019) para implmentação dos algoritmos usaram os dados da *Epileptic Seizure Recognition Data Set*. O conjunto de dados é composto por leituras sinais de EEG de 500 (quinhentos) pacientes durante coletados durante 23,5 segundos. A partir dos sinais geram identificados 5 (cinco) tipos de atividades cerebrais, sendo elas: convulsão, área de tumor, área saudável, olhos fechados e olhos abertos. A partir dos sinais coletados, os algoritmos de aprendizagem de máquina foram utilizados para realizar a predição dos cinco estados distintos. Os autores do artigo não disponibilizaram o código fonte implemen-

tado, sendo necessário desenvolvê-lo para comparação entre a acurácia apresentada no artigo e encontrado na implementação.

III. IMPLEMENTAÇÃO DO ARTIGO

Devido a disponibilidade de apresentação do código fonte, foi escolhido o *Google Collaboratory* para implementação, dessa forma o professor da disciplina e os alunos puderam executar e testar o código durante a apresentação. O código está disponível em http://colab.research.google.com/drive/1ev5XQd6kRNvdIhpKL22_XqVXN9cIipcPscrollTo=YgGvZ-fBwfBV.

A. Importação das bibliotecas

Inicialmente foram importadas as bibliotecas *pandas*, *numpy*, *matplotlib* e *sklearn*. Desta ultima utilizou-se os pacotes *train_test_split*, *LogisticRegression*, *KNeighborsClassifier* e *metrics*. Para armazenamento dos dados foi utilizado o módulo *drive* do *google.colab*, assim como o código fonte produzido nas atividades contextualizadas. Utilizando o princípio da modularização, foi criada uma biblioteca, que armazenada no *google drive*, possibilitou a reutilização dos códigos desenvolvidos ao longo da disciplina. O usuário poderá escolher o nome do arquivo em que desejará salvar os resultados gerados.

B. Exploração dos dados

Após importação das bibliotecas, é utilizado o *pandas* para leitura do arquivo *csv* armazenado no *Google Drive*. Nesse momento, os dados são explorado mediante o uso o *pandas.head()*, *pandas.info()* e *pandas.columns*. No último comando é possível identificar os *indexs* dos dados. Utilizando os conhecimentos aprendidos na disciplina, efetuou-se o acesso as colunas desejadas, separando-as em *data* e *target*. A primeira consiste nos sinais de EEG, enquanto a segunda apresenta estado do paciente, descritos na Seção II. Randomicamente foi escolhido um conjunto de dados para cada estado e plotado os sinais utilizando a biblioteca *matplotlib*. No *plot* dos gráficos foi utilizada uma estrutura de repetição para percorrer os cinco estados distintos.

C. Predição

Os dados foram separados em dois conjuntos: dados de treino e teste, com uma amostragem aleatória. Os dados de teste correspondem a 20% do espaço amostral. Com

os dados de treino, utilizando o módulo `LogisticRegression` foi realizado o aprendizado do algoritmo, da mesma forma como foi aplicado o `KNeighborsClassifier`, com distância de classificação igual a 1. Após o aprendizado foi realizada a predição utilizando os dados de teste separados anteriormente.

D. Métricas

Da biblioteca *sklearn* foi importada a *metrics*. Por meio dela foi possível medir a acurácia dos resultados. A maior acurácia foi do KNN, com 0.545, valor superior ao modelo da Regressão Logística que ficou com uma acurácia de 0.218. Os resultados foram armazenados em um arquivo texto no *Google Drive*. Para distinguir os diferentes experimentos foi importada a biblioteca *time* e armazenado o a data e horário do experimento, dessa forma é possível criar um histórico das diferentes acurácias alcançadas.

IV. CONSIDERAÇÕES FINAIS

No artigo escolhido, foi observado que os autores tiveram como objetivo reconhecer a crise epilética em um intervalo de tempo de um segundo. Contudo, faz-se necessário uma análise clínica que possibilite a identificação da atividade convulsiva em tempo real. Liu, XI, Zhao e Li (2019) também implementam redes neurais para elevação da acurácia. Porém, devido à limitação técnica temporária esses modelos não foram usados na apresentação. No desenvolvimento da próxima disciplina espera-se que os conhecimentos de inteligência artificial possam ser ampliados e possibilitem maiores aplicações com resultados mais precisos e satisfatórios no projeto em estudo. Almeja-se nos trabalhos futuros a implementação de diferentes algoritmos de aprendizagem de máquina visando o aumento da acurácia alcançada, além da criação de banco de dados com experimentos executados no Instituto Santos Dumont. Espera-se criar um dispositivo de baixo custo que interpretará os dados de EEG do usuário e executará o controle de diferentes equipamentos, sem a necessidade de movimentação dos membros do usuário. Nesse cenário, a disciplina Fundamentos de Programação e Desenvolvimento de Projetos Aplicados a Neuroengenharia, ministrada no ISD, servirá devido às ferramentas de programação oferecidas, necessárias na implementação dos algoritmos de aprendizagem de máquina.

AGRADECIMENTOS

Ao Grande Arquiteto do Universo, pela oportunidade da busca pelo conhecimento justo e perfeito, e aos companheiros de turma, com os quais muito aprendi.

REFERENCES

- [1] LIU, Haotian; XI, Lin; ZHAO, Ying; LI, Zhixiang. Using Deep Learning and Machine Learning to Detect Epileptic Seizure with Electroencephalography (EEG) Data. 2019. Disponível em: <https://arxiv.org/pdf/1910.02544.pdf>. Acesso em: 06 out. 2019.
- [2] Kiral-Kornek, I., Roy, S., Nurse, E., Mashford, B., Karoly, P., Carroll, T., Payne, D., Saha, S., Baldassano, S., O'Brien, T. and Grayden, D., 2018. "Epileptic seizure prediction using big data and deep learning: toward a mobile system". *EBioMedicine*, 27, pp.103-111. Vancouver.
- [3] Antoniadis, A., Spyrou, L., Took, C.C. and Sanei, S., 2016, September. "Deep learning for epileptic intracranial EEG data". In 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP) (pp. 1-6). IEEE.

- [4] Thodoroff, P., Pineau, J. and Lim, A., 2016, December. "Learning robust features using deep learning for automatic seizure detection". In *Machine learning for healthcare conference* (pp. 178-190).
- [5] UMLS: UCI Machine Learning Repository. <https://www.nlm.nih.gov/research/umls/index.html>
- [6] Zscheischler, J. et.al (2016, August 8). Short-term favorable weather conditions are an important control of interannual variability in carbon and water fluxes. Retrieved from <https://doi.org/10.1002/2016JG003503>