**Phân loại cơn động kinh bằng cách sử dụng máy học Explainable**

Dựa trên thông tin từ các nguồn được cung cấp, dưới đây là tóm tắt về nghiên cứu phân loại co giật động kinh dựa trên dữ liệu EEG sử dụng thuật toán học máy và trí tuệ nhân tạo giải thích được:

•

Động kinh là một rối loạn thần kinh... được đặc trưng bởi các cơn co giật tái phát1... và ảnh hưởng đến hàng triệu người trên toàn thế giới.... Việc **phân loại chính xác các cơn co giật là rất quan trọng** cho chẩn đoán và điều trị hiệu quả…

•

Các thuật toán học máy (ML) gần đây đã cho thấy tiềm năng lớn trong việc phân tích các bản ghi EEG (điện não đồ) và phân loại các cơn co giật động kinh.... Tuy nhiên, bản chất "hộp đen" của các thuật toán này có thể gây khó khăn trong việc hiểu cách chúng đưa ra quyết định....

•

Nghiên cứu này tập trung vào việc **phân loại co giật động kinh sử dụng các thuật toán học máy và trí tuệ nhân tạo giải thích được (Explainable AI - XAI)**....

•

Mục tiêu là phát triển một hệ thống dựa trên máy tính để phân loại chính xác các cơn co giật động kinh dựa trên bản ghi EEG. và khám phá việc sử dụng các thuật toán học máy khác nhau cho mục đích này…. Nghiên cứu cũng nhằm mục đích chứng minh tiềm năng của thuật toán học máy và XAI trong việc phân loại chính xác co giật động kinh và nâng cao sự hiểu biết về các mô hình này….

•

Nhóm nghiên cứu đã sử dụng **một tập dữ liệu EEG có sẵn công khai** từ 500 cá nhân.... Mỗi bản ghi bao gồm 4097 điểm dữ liệu được lấy mẫu trong 23.6 giây.... Dữ liệu được xử lý trước bằng cách chia thành 23 đoạn, mỗi đoạn 178 điểm dữ liệu....

•

Các nhãn được gán cho mỗi bản ghi dựa trên việc cá nhân đó có cơn co giật động kinh (lớp 1) hay không (các lớp 2-5).... Mặc dù có năm lớp trong tập dữ liệu gốc, nghiên cứu này đã thực hiện phân loại nhị phân bằng cách xem xét lớp 1 đối với các lớp còn lại....

•

Tập dữ liệu gốc bị **mất cân bằng** với 9200 mẫu không co giật và 2300 mẫu co giật12. Để cân bằng tập dữ liệu, phương pháp lấy mẫu tăng (up-sampling) đã được triển khai bằng cách sử dụng "SMOTE" + "ENN"...

•

Dữ liệu đã được chia thành ba tập: tập huấn luyện, tập kiểm tra và tập xác thực theo tỷ lệ **60:20:20**…

•

Nghiên cứu đã đánh giá hiệu suất của nhiều thuật toán phân loại khác nhau9, bao gồm Logistic Regression, K-nearest neighbors, Support Vector Machine (SVC), Naive Bayes Classifier, Random Forest Classifier, và Gradient Boosting (XGB)....

•

Các chỉ số hiệu suất như **độ chính xác (accuracy), độ chính xác (precision), độ thu hồi (recall), f1-score, và độ đặc hiệu (specificity)** đã được đo lường cho tất cả các thuật toán bằng cách sử dụng ma trận nhầm lẫn (confusion metrics).... Việc đánh giá bằng các chỉ số này là rất quan trọng để xác định hiệu quả của mô hình trong môi trường lâm sàng.

•

Để giải thích các mô hình và hiểu rõ hơn về quá trình ra quyết định của chúng, các kỹ thuật XAI như **SHAP và LIME** đã được sử dụng.... Việc sử dụng các kỹ thuật XAI như SHAP và LIME trong bối cảnh chẩn đoán động kinh có thể là một cách tiếp cận mới. XAI giúp tăng cường khả năng diễn giải và sự hiểu biết về các mô hình học máy....

•

**Kết quả tốt nhất** về độ chính xác đạt được là **96.1% khi sử dụng thuật toán Random Forest**3.... Các thuật toán khác có độ chính xác là: SVC 95.7%, KNN 93.7%, NB 92.3%, LR 90.1%, và XGB 91.7%17…

•

Thuật toán Random Forest (RF) cũng cho thấy **điểm số cao nhất cho tất cả bốn chỉ số hiệu suất** (precision, recall, F1-score, và specificity), theo sau là thuật toán SVC.... Thuật toán Logistic Regression (LR) có điểm số thấp nhất cho cả bốn chỉ số này.

•

Phân tích SHAP cho thấy **đặc trưng "x32" là đặc trưng quan trọng nhất**, có tác động lớn nhất đến việc đưa ra dự đoán của mô hình...

•

Tóm lại, nghiên cứu đã chứng minh **tiềm năng của thuật toán học máy (đặc biệt là Random Forest) trong việc phân loại chính xác các cơn co giật động kinh**. Việc sử dụng kỹ thuật SHAP và LIME đã cung cấp những hiểu biết có giá trị về quá trình ra quyết định của mô hình và nâng cao khả năng diễn giải của chúng, điều này có thể góp phần vào việc phát triển các phương pháp hiệu quả hơn để chẩn đoán và điều trị động kinh. Công việc này đóng góp vào những nỗ lực không ngừng nhằm cải thiện chẩn đoán và điều trị động kinh sử dụng thuật toán học máy và kỹ thuật trí tuệ nhân tạo giải thích được…