Explaining Biomarker Response to Anticoagulant Therapy in Atrial Fibrillation: A Study of Warfarin and Rivaroxaban with Machine Learning Models

Luan Sena
luanborges@ufmg.br
18/11/2024

DCC
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
UFMG

Contextualization

Atrial Fibrillation is a common form of arrhythmia, affecting approximately 10% of individuals by the age of 80.

Disrupts blood flow, causing stasis and increasing the risk of thromboembolic events such as ischemic stroke and systemic embolism.

To prevent such events anticoagulant therapy is used. Warfarin, a traditional anticoagulant, and direct oral anticoagulants (DOACs), such as rivaroxaban are commonly used medications in AF treatment.

Motivation

Create models capable of identifying if a patient had AF via biomarker data.

Create models capable of identifying the medicine the patients with AF were using.

Learn how different was the biomarker response from the use of different anticoagulants, given a small dataset of patients.

Separate relevant biomarkers in order to learn how they influence the model.

Used Data

195 individuals, including 109 in the control group (without AF) and 86 patients with AF. Among the AF patients, 47 were using warfarin and 39 were using rivaroxaban.

The biomarkers used comprise patient data from characterization, blood count, lipid profile, coagulation, inflammatory and cardiac diseases of the individuals.

Relevant Biomarkers

Some relevant biomarkers in the context of atrial fibrillation are:

- **Inflammatory**: Interleukins (IL-2, IL-4, IL-8, IL-10).
- **Coagulation**: Peak, Microparticles (MPE, MPP), Prothrombin Fragment 1+2, ETP (Endogenous thrombin potential).
- **Lipid Profile**: Total Cholesterol.
- **Characterization**: Dyslipidemia (hypercholesterolemia).
- **Blood Count**: PLT (Platelets).
- **Cardiac Diseases**: sICAM-1 (Soluble Intercellular Adhesion Molecule).

Synthetic Data Generation

To further enhance the visualization and analyses of the importance of different features in model predictions. We amplify our dataset via generated synthetic data.

Synthetic data refers to artificial data generated by algorithms that mimic the statistical properties of real-world data.

We used both statistical methods (Gaussian Copula), as well as generative methods (Tabular Variational Autoencoder).

Model

For building our classifiers we utilized an implementation of the LightGBM (Light Gradient Boosting Machine) algorithm.

Boosting algorithms consist of iteratively learning weak classifiers with respect to a distribution and adding them to a final strong classifier.

LightGBM uses tree based algorithms in order to build predictions.

Results and Ablations

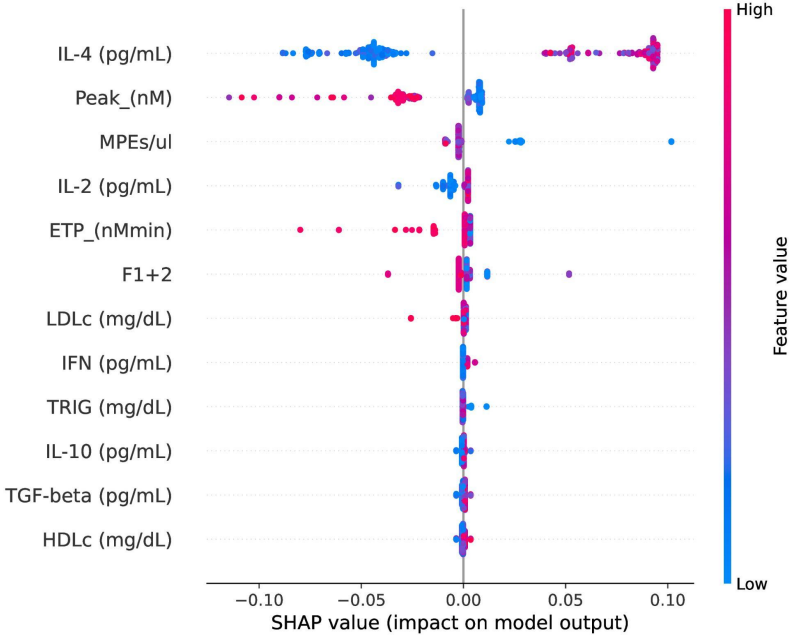
Predicting AF vs Control

Feature Group	ROC AUC Average ± STD
Full Dataset	0.999 ± 0.001
Without Coagulation Biomarkers	0.989 ± 0.017
Without Inflammation Biomarkers	0.981 ± 0.023

Predicting Rivaroxaban vs Warfarin

Feature Group	ROC AUC Average ± STD
Full Dataset	0.986 ± 0.027
Without Coagulation Biomarkers	0.945 ± 0.029
Without Inflammation Biomarkers	0.971 ± 0.036

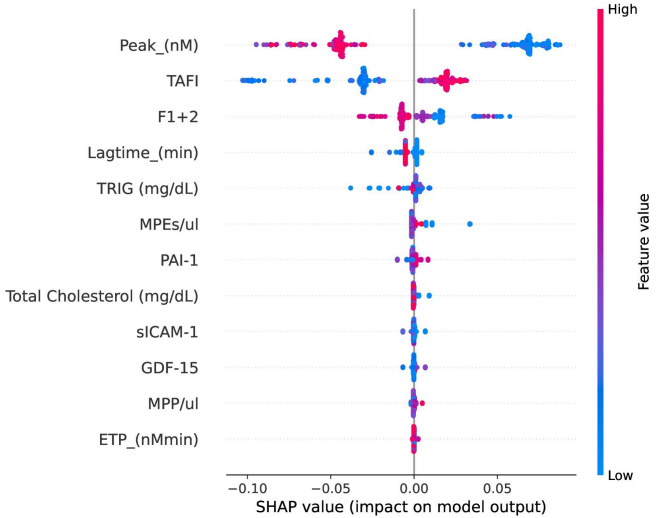
Explainability - Control vs AF - Full Dataset



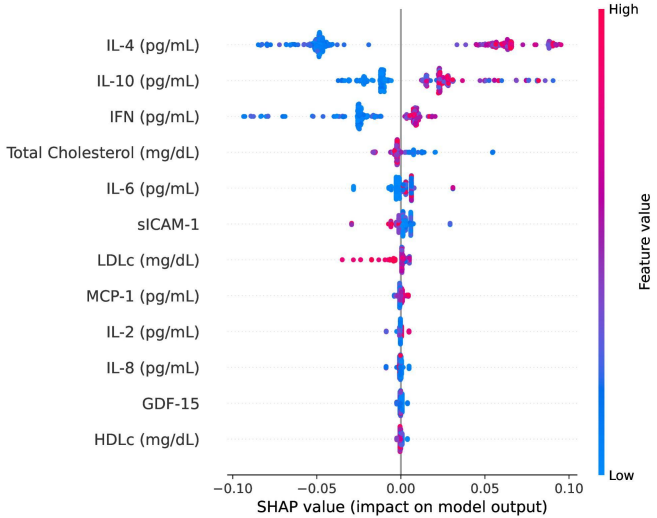
DCC
DEPARTAMENTO DE
CIÊNCIA DA COMPUTAÇÃO
UFMG

Explainability - Control vs AF - Ablations

Without inflammatory markers



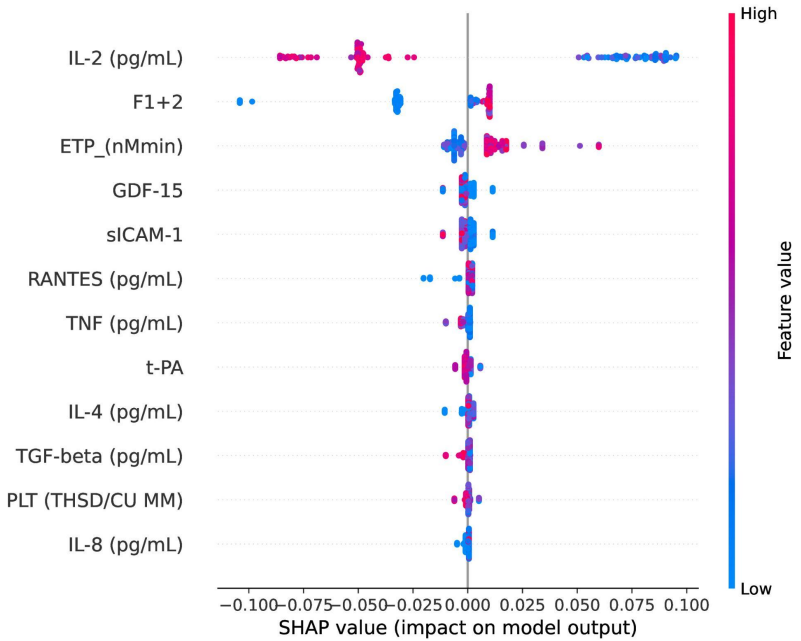
Without coagulation markers



DCC
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

UFMG

Explainability - Warfarin vs Rivaroxaban - Full Dataset

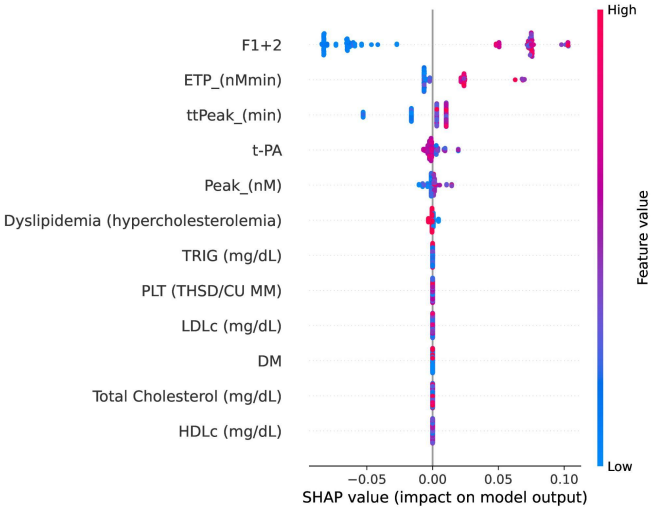


DCC
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

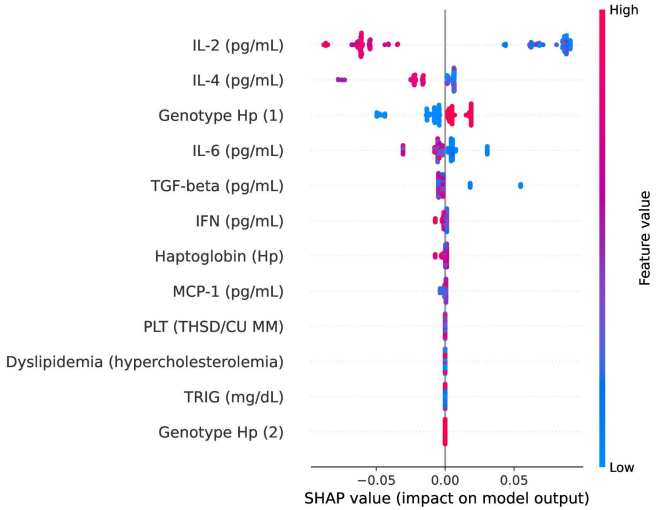
UFMG

Explainability - Warfarin vs Rivaroxaban - Ablations

Without inflammatory markers



Without coagulation markers



DCC
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

UFMG

Conclusion

Some contributions of our work were:

- We were able to build high precision models for predicting AF, and medication used on patients using machine learning models.
- Via explainability we showed which parameters were influencing patients using different medications.
- Explored the area of atrial fibrillation in order to further research how machine learning can be used to improve the prediction and treatment of this condition.

Explaining Biomarker Response to Anticoagulant Therapy in Atrial Fibrillation: A Study of Warfarin and Rivaroxaban with Machine Learning Models

Luan Sena
luanborges@ufmg.br
18/11/2024

DCC
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
UFMG