



# Nasa Space Challenge - Develop the Oracle of DSCOVR

AstroNet: Ariel Carvalho Alves

# Motivação do Projeto

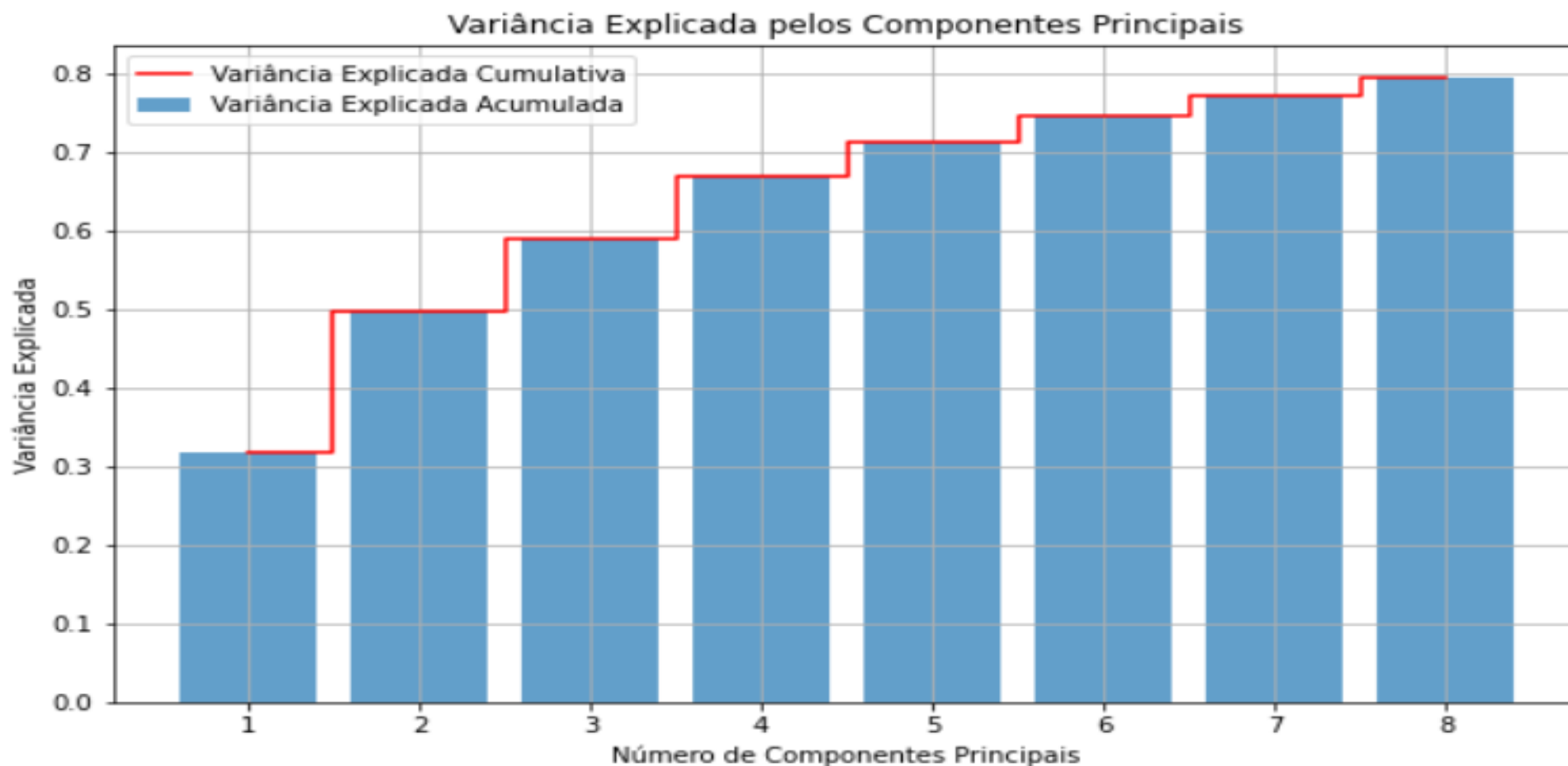
- Previsão de tempestade solar a partir da medição do magnetismo e do vento (dados extraídos do DSCOVR)
- Utilização de Deep Learning para identificar anomalias
- As previsões podem ajudar a identificar quando ocorrerão interrupções nas redes elétricas, satélites, aviação, GPS e telecomunicações

# Tratamento dos Dados

- Tratamento de valores ausentes com descarte de algumas colunas e substituição pela mediana
- Utilização da técnica PCA para reduzir as dimensões depois de combinar mais de 50 colunas
- Criação da variável target binária Anomalia para identificar outliers (valores extremos) a partir do principal componente que melhor representa a variação nos dados.

# Tratamento dos Dados

- Foram definidos 8 componentes principais para representar mais de 80% dos dados de forma significativa

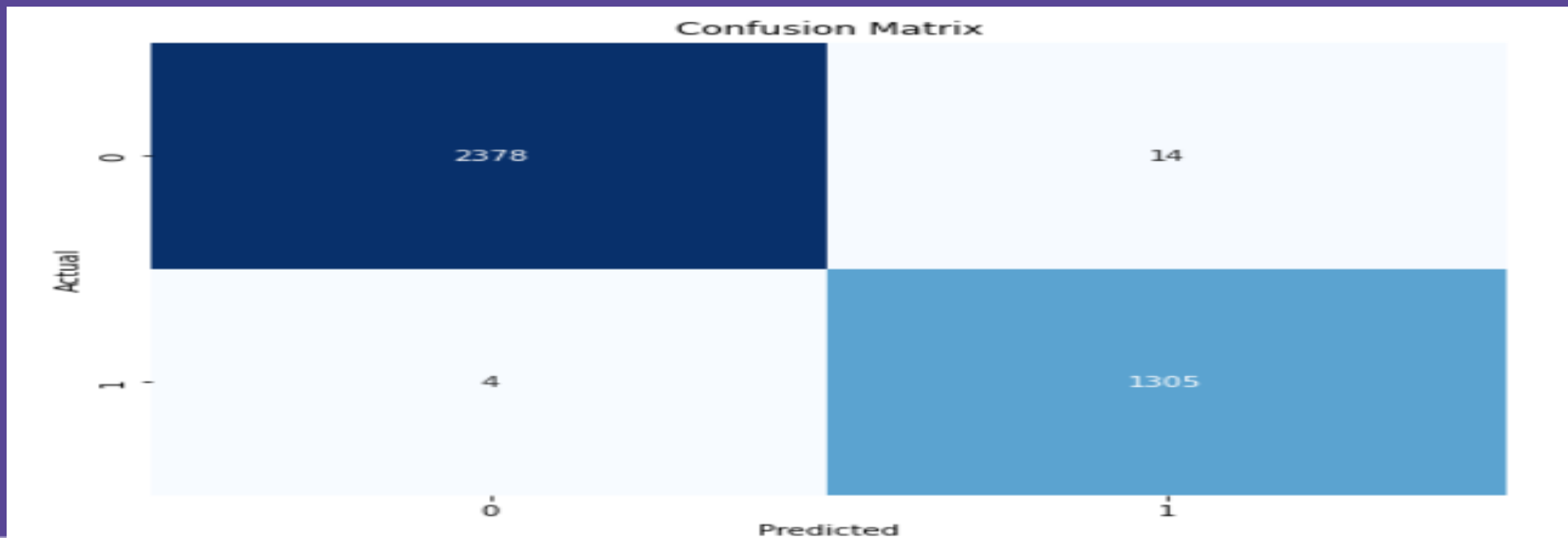


# Construção do Modelo

- Utilização do Tensorflow
- Otimizador Adam para descida do gradiente estocástico ajustado de forma dinâmica
- Binary Crossentropy para verificar a entropia cruzada e medir a divergência entre a distribuição da probabilidade prevista e os dados reais
- Função de Ativação Sigmoid para binarizar a classificação (anomalia ou não)

# Métricas e Melhorias

- Recall foi escolhido para avaliar a sensibilidade do modelo e o F1-Score para avaliar o impacto (tendo em vista que é a média harmônica entre o recall e a accuracy)
- Há espaço para melhorias e redução de overfitting por parte do modelo tendo em vista que as métricas utilizadas atingiram valores elevados demais



**Obrigado pela disponibilidade!**

