

# RELATÓRIO

Reconhecimento de datolologia usando machine learning



## O QUE É DATILOLOGIA?

Os alfabetos manuais são recursos muito importantes das línguas de sinais, uma vez que representam as letras do alfabeto escrito de um determinado idioma. Cada letra é representada com uma configuração de mão, e podem apresentar movimento ou não.

Essa representação é muito importante uma vez que possibilita a soletração de palavras, o que cria uma interlocução entre a língua verbal e a língua visual.

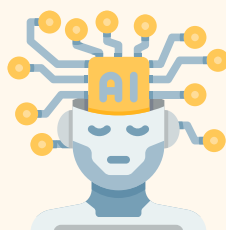
## MEMBROS DO GRUPO:

Gabriela Cristina Schmitt  
Jennifer M. de P. Goberski  
Vinicius Dionizio Patrocinio

Data: 02 de junho

## OBJETIVO:

Com base no alfabeto manual da LIBRAS, queremos criar um modelo que reconheça as configurações de mãos apresentadas por vídeo e consiga inferir a qual letra essa configuração representa.



## IMPORTANTE:

O nosso modelo foi treinado para reconhecer letras estáticas, então as letras H, J, K, X e Z não foram incluídas e não podem ser inferidas! Além disso nosso dataset contém imagens destras e canhotas dos sinais.

## MATERIAIS:

- Notebook: Colab Google
- Programas: Anaconda e VsCode

## ETAPAS:

- Normalização
- Balanceamento
- Hyperparametrização
- Random Forest
- Cross Validation

## DATASET:

- 21 letras mapeadas
- 3600 imagens para cada letra



## PROCEDIMENTO:

### Extração de dados

1. Criação do DataBase com imagens nossas e extração das coordenadas de cada um dos vinte vetores mapeados nas mãos (Hand Landmarks);
2. Serializar em um csv as coordenadas e qual letra cada imagem representa;
3. Ler o csv e gerar um Dataframe com os dados contidos nele;

### Preprocessing

4. Separar os atributos das classes da base;
5. Normalizar e balancear ambos;

### Treinamento

6. Criar uma RandomForest e hiperparametriza-la;
7. Criar uma árvore de decisão e usá-la para alimentar a RandomForest com os atributos e classes balanceados e normalizados;
8. Usar cross validation com a árvore treinada e com os dados para termos as métricas;
9. Gerar a matriz de contingência;

### Inferência

10. Conectar a webcam e processar a inferência frame a frame, mostrando o resultado da classificação na tela.

## RESULTADOS:

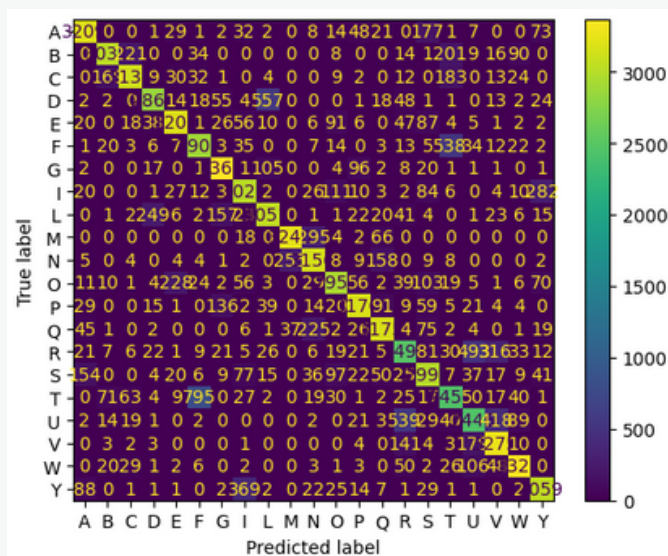
Para conseguirmos mensurar os resultados da rede, foi utilizada a técnica de cross validation na base normalizada e balanceada. Através dela, podemos extrair as seguintes métricas:

Sensibility: 83.42%

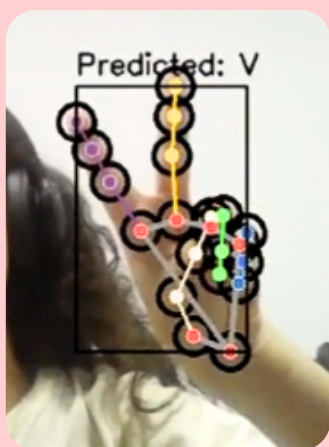
Specificity: 84.15%

Accuracy: 83.46%

Além disso, também geramos a matriz de contingência com as classes inferidas pelo modelo e com as classes reais associadas aos dados utilizados.



## CONCLUSÃO:



A inferência teve resultados positivos e, em geral, prediz e exibe no console a letra correta sinalizada ao vivo pela webcam. Porém, ocorrem erros em uma proporção considerável, além de ocorrerem erros relacionados a distância da mão em relação à câmera, uma vez que se torna mais difícil a extração das coordenadas quando a mão está mais distante ou está um pouco inclinada.

