APS – 6º semestre - 2018

DESENVOLVIMENTO DE UM SISTEMA BIOMÉTRICO CAPAZ DE IDENTIFICAR ESPÉCIES DE PLANTAS A PARTIR DE SUAS FOLHAS

UNIP – Universidade Paulista

CC6P52

Integrantes:

* Carlos Eduardo de Castro – N902JC3
* Manoel de Freitas Gouvêa Junior – B999AI0
* Marcos Guilherme Afonso de Paula – N953BD7

Sumário

[Introdução, objetivo e motivação do trabalho 2](#_Toc529889629)

[Principais Técnicas Biométricas 4](#_Toc529889630)

[Fundamentos dos classificadores 4](#_Toc529889631)

[kNN 4](#_Toc529889632)

[Sobre 4](#_Toc529889633)

[Fundamentação Matemática 4](#_Toc529889634)

[Funcionamento 4](#_Toc529889635)

[MLP (Multi Layer Perceptron) 4](#_Toc529889636)

[Sobre 4](#_Toc529889637)

[Fundamentação Matemática 4](#_Toc529889638)

[Funcionamento 4](#_Toc529889639)

[Regressão Logística 4](#_Toc529889640)

[Sobre 4](#_Toc529889641)

[Fundamentação Matemática 4](#_Toc529889642)

[Funcionamento 4](#_Toc529889643)

[Planejamento e experimentação 4](#_Toc529889644)

[Condução dos testes e Tabulação de Resultados 4](#_Toc529889645)

[kNN 4](#_Toc529889646)

[Desempenho e Avaliação de resultados 4](#_Toc529889647)

[MLP (Multi Layer Perceptron) 4](#_Toc529889648)

[Desempenho e Avaliação de resultados 4](#_Toc529889649)

[Regressão Logística 4](#_Toc529889650)

[Desempenho e Avaliação de resultados 4](#_Toc529889651)

[Conclusão e Discussão dos Resultados 5](#_Toc529889652)

[Referências Bibliográficas 5](#_Toc529889653)

[Apêndice A 5](#_Toc529889654)

# Introdução, objetivo e motivação do trabalho

A história da Inteligência artificial teve início logo após a segunda guerra mundial com Alan Turing, conhecido por ser o pai da computação e um dos pioneiros do campo da IA, em 1956 em uma conferência no campus do Darthmouth College foi fundado o campo de pesquisa em inteligência artificial, definido como “A ciência e engenharia de produzir máquinas inteligentes”.

Porém como tudo que se é feito estritamente em campo teórico existiam problemas para implementação prática das ideias, com a inteligência artificial não foi diferente, na época de sua idealização não existia processamento suficiente para a aplicação dos conceitos, somente em 1997 a IBM® construiu um computador capaz de realizar tal nível de processamento, ele foi batizado como Deep Blue e foi o computador responsável por derrotar o enxadrista Garry Kasparov, tido como o melhor jogador de xadrez de toda a história, apesar de que o Deep Blue fosse capaz de realizar o processamento de uma IA ainda estava muito longe do que conhecemos nos dias atuais.

Indo para um momento mais recente na história, em 2011 a mesma empresa responsável pelo Deep Blue, a IBM®, pois a prova sua nova inteligência artificial, o Watson, em um jogo de perguntas e respostas chamado Jeopardy, mesmo competindo com os maiores vencedores desse jogo o supercomputador se mostrou capaz de dominar todas as rodadas do game.

Claro que a Inteligência Artificial não se limita a apenas a jogos e diversão é um campo que demonstra um grande avanço nas mais variadas áreas, podendo ser aplicada aos mais diferentes tipos de problemas, mas também ela não é capaz de resolver tudo ainda estamos em um momento de evolução onde certos problemas ainda estão fora do alcance dessa área tão promissora.

Conforme comentado anteriormente a área de inteligência artificial é muito ampla e possui tendências de crescer cada vez mais conforme os avanços científicos ocorrem, nas próximas linhas iremos comentar brevemente sobre algumas dessas áreas.

Uma das áreas mais antigas dentro de IA é o campo das buscas, que é focada em localizar soluções, respostas, em grafos de formas cada vez mais eficientes, outra área bastante famosa é a área de Machine Learning ou aprendizado de máquina, que consiste em criar algoritmos com a capacidade de aprender padrões para então realizar predições com base no conhecimento adquirido, de forma automática, outra área a citarmos é a área de PLN ou Processamento de Linguagem Natural, essa é uma área que consiste em analisar e produzir textos se baseando nas línguas faladas pelos seres humanos, como o português ou então o inglês, continuar nessas citações levaria muito tempo e seria o suficiente para uma monografia apenas para isso.

Para esse projeto utilizaremos algumas das ferramentas fornecidas pela área de Machine Learning, mais especificamente falando utilizaremos ferramentas para a solução de problemas de classificação, que são problemas cuja natureza é determinar a categoria de um certo individuo informado ao programa, um detalhe fundamental é que essas categorias já são conhecidas pelo algoritmo que irá avalia-las, e consiste em um aprendizado supervisionado, isto é, os programadores devem oferecer as respostas corretas para o algoritmo de treinamento para que então ele de forma automática extraia o conhecimento dos dados.

Indo especificamente para nosso projeto o problema consistem em classificar plantas com base em parâmetros retirados de imagens das folhas dessas plantas, nós não realizamos o tratamento dessas imagens tampouco coletamos os dados apenas utilizamos o dataset que já existe e está disponibilizado no site UCI, que é um repositório de detasets muito utilizado para a área de Machine Learning. Os dados originais foram coletados por Rubim Almeida da Silva e o dataset foi criado por Pedro F. B. Silva e André S. Marçal, todos da Universidade do Porto, Portugal.

Essa atividade consiste em realizar o treinamento e avaliação de resultados de três diferentes Inteligências Artificiais trabalhando como classificadores das informações retiradas do dataset. Os detalhes de implementação poderão ser dispensados, inclusive optamos pela utilização do software Orange®, que basicamente consiste em uma interface gráfica para realização de treinamentos, avaliação, montagem de gráficos de dispersão, geração de matrizes de confusão, entre outras opções contando com várias implementações de diferentes IA’s, como motor para tudo isso o Orange® utiliza como núcleo a linguagem python.

Neste projeto os integrantes também aproveitaram a oportunidade para utilizar ferramentas diferentes das até então conhecidas pelo grupo, além do software Orange® aproveitamos para aprimorar os conhecimentos dos integrantes na utilização da plataforma GitHub®, que é um repositório para divulgação de projetos, em sua maioria open-source. Link para o projeto: [Repositório Projeto](https://github.com/ManoelGouveaJunior/APS-SII).

# Principais Técnicas Biométricas

# Fundamentos dos classificadores

## kNN

### Sobre

### Fundamentação Matemática

### Funcionamento

## MLP (Multi Layer Perceptron)

### Sobre

### Fundamentação Matemática

### Funcionamento

## Regressão Logística

### Sobre

### Fundamentação Matemática

### Funcionamento

# Planejamento e experimentação

# Condução dos testes e Tabulação de Resultados

## kNN

### Desempenho e Avaliação de resultados

## MLP (Multi Layer Perceptron)

### Desempenho e Avaliação de resultados

## Regressão Logística

### Desempenho e Avaliação de resultados

# Conclusão e Discussão dos Resultados

# Referências Bibliográficas

# Apêndice A