



UNIVERSIDADE PAULISTA

UNIP – Universidade Paulista

Ciência da Computação (CC)

Orientações para as Atividades Práticas Supervisionadas 5º/6º semestres (2018)

- TEMA
- PROPOSTA DO TRABALHO
- APRESENTAÇÃO DO TRABALHO

Atividades Práticas Supervisionadas – 5º/6º SEMESTRES – CC

I. TEMA:

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

II. PROPOSTA DO TRABALHO

Neste semestre as Atividades Práticas Supervisionadas têm o propósito de explorar conceitos e técnicas ligados ao processamento de imagens, manipulação de informações biométricas, extração de características e a aplicação de técnicas de classificação, típicas da área de aprendizado de máquina, para o reconhecimento e identificação de plantas.

O objetivo do trabalho é realizar, na perspectiva de projetista, a modelagem, treinamento e construção de um sistema inteligente capaz de identificar espécies de plantas a partir de imagens de suas folhas.

A utilização informações reais em problemas de classificação é sempre desejável, uma vez que permite uma análise mais consistente de desempenho e maior credibilidade aos resultados. Assim, o uso de dados reais é bastante importante para a validação da técnica que será adotada pelo grupo. Entretanto, como o foco principal não é a aquisição ou pré-processamento de imagens, os integrantes utilizarão uma base de dados pública que forneça imagens ou dados sobre diferentes espécies de plantas.

Recomendamos o uso do *dataset* LEAF que é disponibilizado no repositório da *UCI Machine Learning* (<https://archive.ics.uci.edu/ml/datasets/leaf>), o qual contém informações sobre 40 espécies diferentes de plantas e cerca de 10 imagens reais para cada uma delas.

Na perspectiva de um projetista de sistemas inteligentes é fundamental analisar o problema a fim de escolher quais técnicas podem ser aplicadas e, após um processo experimental de treinamento e testes destas técnicas, escolher a que oferece melhor desempenho naquela situação.

Pede-se aos alunos que escolham três métodos diferentes de classificação a ser aplicados ao problema proposto, sendo preferencialmente um de cada paradigma: baseado em instâncias, probabilístico, conexionista ou simbólico.

Observações:

- 1) No caso da escolha por uma base de dados de imagens, cada uma das entradas deverá ser pré-processadas de modo a produzir um vetor com características que representem aquela instância. A quantidade de atributos e a forma/métrica para sua obtenção ficará a livre escolha do grupo, desde que atendam as especificações deste trabalho.
- 2) O grupo deverá fazer uma dissertação sobre os conceitos que envolvem a aquisição automática de conhecimento (aprendizado de máquina) e a sua aplicação em problemas de classificação. Deve ainda explicar os fundamentos para cada uma das técnicas escolhidas pelo grupo e de que forma ela estrutura o conhecimento.
- 3) É permitido o uso de softwares específicos, frameworks, bibliotecas de *machine learning* ou de planilhas eletrônicas no desenvolvimento do trabalho uma vez que esta APS tem por objetivo avaliar a capacidade do grupo em modelar e desenvolver um método para a classificação de plantas.
- 4) O claro encadeamento de ideias, etapas do processo e consistência tanto da modelagem quanto da execução serão objetos de avaliação neste trabalho, assim como a interpretação dos resultados e as conclusões apresentadas pelo grupo.
- 5) Nas considerações finais, espera-se que o grupo apresente suas reflexões sobre o efeito desse trabalho na sua formação e a interdisciplinaridade envolvida no mesmo.
- 6) O nível de refinamento, funcionalidade extras e análises de desempenho entre diferentes técnicas de classificação terão impacto direto na nota final deste trabalho.
- 7) A nota atribuída ao trabalho entregue configurará na nota de Atividades Práticas Supervisionadas - APS

III. APRESENTAÇÃO DO TRABALHO

1. O grupo deverá ser composto por, no máximo, 4 alunos. A formação de um grupo com um número menor do que 3 dependerá de aprovação do professor responsável pela atividade.
2. Todas as etapas do trabalho deverão ser escritas em fonte ARIAL 12, espaçamento 1,5, margem direita 2,5 cm e margem esquerda 2,5 cm. O trabalho deverá ter formato A4, encadernado (espiral) com capa transparente.

3. Limites de páginas

Introdução, objetivo e motivação: de 2 a 3 páginas no máximo.

Fundamentos das técnicas biométricas (conceitos gerais) e extração de características a partir de imagens: de 3 a 4 páginas no máximo.

Fundamentos dos classificadores escolhidos: de 4 a 8 páginas no máximo.

Planejamento dos experimentos, condução dos testes e tabulação dos resultados: mínimo de 3 páginas e máximo de 6 páginas.

Discussão dos resultados e análise: mínimo de 3 páginas e máximo de 5 páginas.

4. Estrutura do trabalho:

1. Capa: identificando o curso, o tema, a relação de alunos do grupo (nome/RA)
2. Índice
3. Introdução
4. Objetivo e motivação do trabalho
5. Fundamentos das principais técnicas biométricas (conceitos gerais)
6. Extração de características a partir de imagens
7. Aprendizado de máquina e classificação
8. Relatório sobre o processo experimental
9. Resultados, análise e conclusões.
10. Bibliografia
11. Ficha de Atividades Práticas Supervisionadas (individual), a ser digitalizada e anexada ao fim do documento.