

Universidade de Brasília
Departamento de Ciência da Computação
Projeto 2, Introdução à Inteligência Artificial, Turma A, 2/2020
Prof. Díbio

Neste projeto serão estudados e avaliados aspectos de aplicação de modelos de árvores de decisão, e modelos de florestas randômicas. Desde início de 2020, infelizmente, o mundo tem sido atingido pela pandemia de covid-19. Dados acerca de como diagnosticar precisamente pacientes, através de exames de sangue rotineiros em hospitais, e possíveis prognósticos de gravidade da doença (e.g. necessidades de UTI), têm sido disponibilizados. Neste projeto você deverá usar os dados editados e disponibilizados no arquivo anexo dataset.zip

e realizar em Python, um projeto, com as seguintes características:

- a) Três (3) pontos para implementação modular, comentada, de algoritmo C4.5 (árvore de decisão), “Random Forest”, e qualquer outro classificador a sua escolha, sem erros de compilação e que recebe como entrada os dados como indicados acima (USP);
- b) Dois (2) pontos para realizar uma análise exploratória dos dados, limpando e gerando variáveis e tabelas relacionadas (e.g. <https://towardsdatascience.com/exploratory-data-analysis-visualization-and-prediction-model-in-python-241b954e1731>);
- c) Três (3) pontos para otimização dos parâmetros da “Random Forest”;
- d) Dois (2) pontos para os valores finais indicando os melhores diagnósticos possíveis dos dados disponibilizados.

O projeto deve ser feito e entregue de forma individual, em um arquivo compactado, e com todos os elementos de comentários em um arquivo tipo *readme*. O projeto deve ser entregue somente via aprender3.unb.br, seguindo os prazos estipulados.