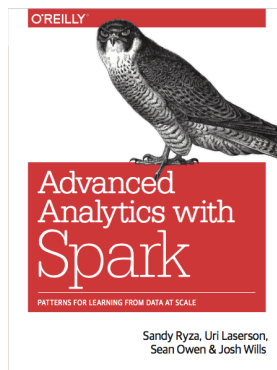


Trabalho #4 – Projeto Final

Neste projeto final o seu grupo deverá utilizar como base o seguinte livro texto, disponível no SIGAA. Este livro apresenta diversos projetos de ciência de dados usando o framework Spark.



Cada grupo será responsável por um projeto específico deste livro, descrito em um capítulo, e deverá executar as seguintes tarefas:

1. Implementar em linguagem Python o projeto descrito no capítulo correspondente ao seu projeto;
2. Criar um notebook PySpark com a solução e submeter via SIGAA;
3. Avaliar e implementar possíveis melhorias ao projeto proposto;
4. Realizar apresentação oral sobre a sua implementação, onde o seu grupo deverá explicar a implementação e as melhorias implementadas.

O seu projeto será avaliado da seguinte forma:

1. (6,0 pontos) Implementação em Python + Spark do projeto estudado;
2. (3,0 pontos) Apresentação oral do projeto implementado;
3. (1,0 pontos) Melhoria implementada;

OBSERVAÇÃO IMPORTANTE:

Para os grupos que forem selecionados para implementar o Capítulo 4 do referido livro, visto que o mesmo apresenta uma implementação simples, deverão ainda: (1) Melhorar a acurácia dos algoritmos apresentados utilizando passos de pré-processamento, por exemplo, seleção de features, categorização, normalização, etc e (2) Comparar os resultados dos algoritmos encontrados no livro com o resultado da aplicação com algoritmo de Redes Neurais.