<http://www.eduardosan.com/2015/06/09/extraindo-dados-de-redes-sociais-com-python/>

<http://pythonclub.com.br/extraindo-texto-de-imagens-com-python.html>

<https://blog.codeexpertslearning.com.br/lendo-imagens-uma-abordagem-%C3%A0-ocr-com-google-tesseract-e-python-ee8e8009f2ab>

<https://github.com/tesseract-ocr/tesseract/wiki>

<https://stackoverflow.com/questions/42831662/python-install-tesseract-for-windows-7>

**duas a três páginas**.

Uma proposta de projeto final inclui sete pontos principais. São eles:

* **O domínio do background** do projeto - ou seja, o domínio do campo de pesquisa de onde o projeto é originado;

O estudante deve detalhar brevemente as informações do background no qual o projeto é proposto. Informações históricas relevantes ao projeto devem ser incluídas. Deve estar claro como e por qual motivo o problema do domínio pode ou deve ser resolvido. Pesquisas acadêmicas relacionadas devem ser apropriadamente mencionadas. Encorajamos que seja mencionada qual foi a motivação pessoal do estudante para que ele realizasse a investigação de determinado problema nesse domínio, porém, isso não é obrigatório.

* **A enunciação do problema** - a definição clara do problema que será investigado, permitindo o surgimento de sua solução;

O estudante deve claramente descrever o problema que será resolvido. O problema deve ser bem descrito e ter ao menos uma solução possível e relevante. Além disso, o problema deve ser quantificável, mensurável e replicável.

* **Os conjuntos de dados e os inputs** - os dados ou inputs usados para resolver o problema;

O(s) conjunto(s) de dados e input(s) a serem usados no projeto devem ser completamente descritos. Devem ser incluídas as informações de como o conjunto de dados ou input são(foram) obtidos, juntamente com suas características. Deve ser nítido como o(s) conjunto(s) de dados e o input(s) serão usados no projeto e se seu uso é apropriado, dado o contexto do problema.

* **A enunciação da solução** - a exposição da solução para o problema apresentado;

O aluno deve descrever claramente a solução do problema. A solução deve ser relevante ao domínio do projeto e apropriada para o(s) conjunto(s) ou input(s) fornecidos. Além disso, deve ser quantificável, mensurável e replicável.

* **Os modelos de benchmark** - alguns exemplos simples ou históricos de resultados que permitam comparar a solução encontrada por você;

Um modelo de benchmark relacionado ao domínio é apresentado, assim como a enunciação do problema e a proposição de uma solução. Idealmente, o modelo de benchmark criado pelo aluno fornece o contexto para métodos existentes ou informações conhecidas no domínio e no problema apresentado, que então podem ser objetivamente comparados com a solução indicada pelo aluno. O modelo de benchmark é claramente definido e mensurável.

* **Conjunto de métricas de avaliação** - representações funcionais de como a solução pode ser mensurada;

O estudante propôs ao menos uma métrica de avaliação que pode ser usada para quantificar a performance tanto do modelo de benchmark como do modelo se solução apresentada. A(s) métrica(s) de avaliação(ões) proposta(s) é(são) apropriadamente relacionada(s) ao contexto dos dados, a declaração do problema e a solução pretendida.

* **Esboço do design do projeto**- como a solução será desenvolvida e como os resultados serão obtidos.

O aluno sintetizou um fluxo de trabalho teórico que permitiu chegar no problema exposto. Foi feita uma discussão sobre quais estratégias poderiam ser empregadas, sobre como a análise dos dados poderia ter sido solicitada, ou sobre quais algoritmos poderiam ser considerados. O workflow e a discussão proposta estão alinhados com as qualidades do projeto. Pequenas visualizações, pseudocódigos ou diagramas são encorajados, mas não são obrigatórios.

* **Apresentação**

A proposta segue uma estrutura bem organizada e será rapidamente entendida pelo seu público-alvo. Cada seção é escrita de maneira clara, concisa e de uma maneira específica. Poucos erros gramaticais e ortográficos estão presentes. Todas as fontes usadas foram mencionadas e corretamente citadas.

Pense, primeiramente, em uma área técnica ou alguma área pela qual seja apaixonado, como a robótica, a realidade virtual, as finanças, o processamento de linguagem natural ou mesmo a inteligência artificial. As possibilidades são infinitas! Então, escolha um problema existente dentro desse campo em que está interessado e no qual você poderia solucionar um problema ao utilizar os algoritmos e técnicas de machine learning. Certifique-se de coletar todas as fontes necessárias (como as bases de dados, inputs e a própria pesquisa) para concluir este projeto e de fazer as citações apropriadas sempre que necessário.