

FIAP

NBA



NLP

Processamento de Linguagem Natural

Prof. Anderson Dourado

Contatos

profanderson.dourado@fiap.com.br



/anderson-dourado

FIAP MBA⁺

Copyright © 2023 | Professor Anderson Vieira Dourado
Todos os direitos reservados. Reprodução ou divulgação total ou parcial deste documento, é expressamente proibido sem consentimento formal, por escrito, do professor/autor.

Aula 1

Introdução e Técnicas de
Pré-processamento

Aula 2

Modelos de classificação
(Análise de Sentimentos)

Aula 3

Mineração de textos e
Modelagem de tópicos

Aula 4

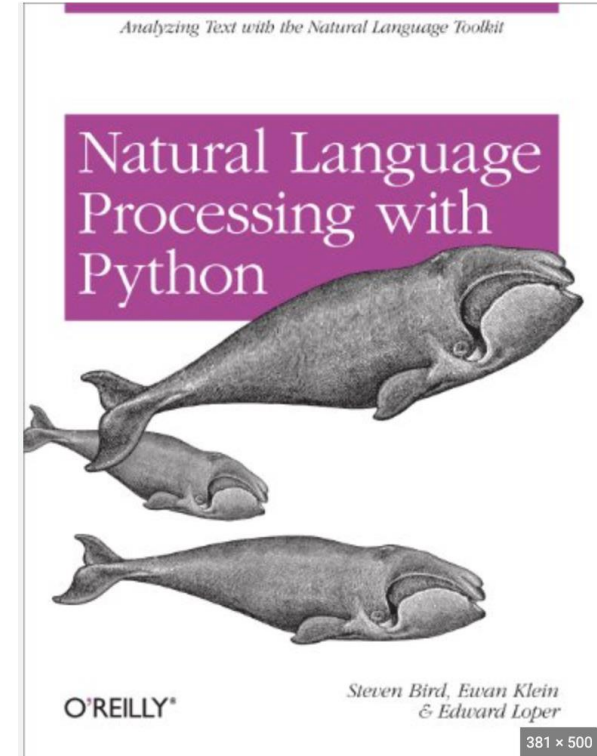
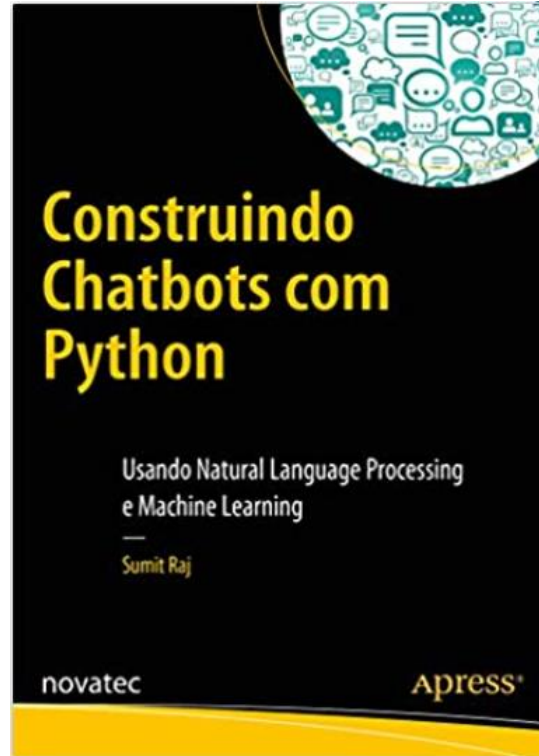
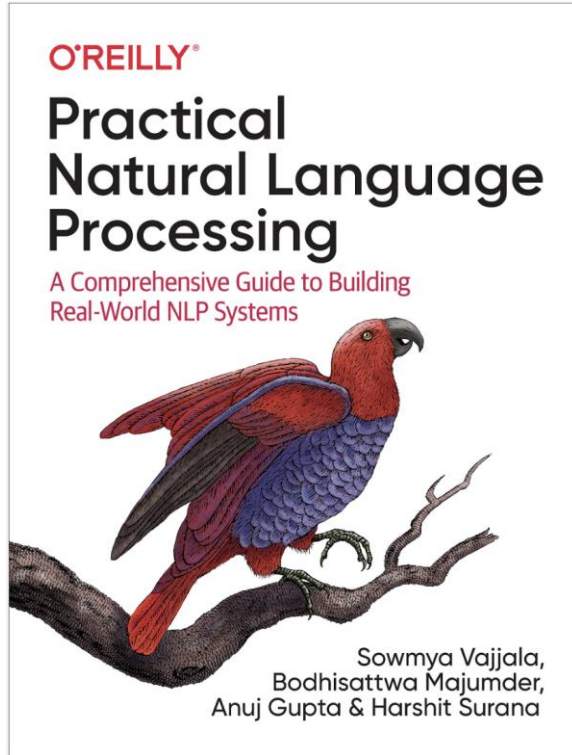
Embeddings e Word2Vec

Aula 5

Revisão +
Generative IA

Aula 6

Generative IA +
Case (classificação)



Todos os códigos usados em sala estarão disponíveis no portal gradativamente.

- Teoria
- Demo = Demonstração prática em aula
- Exercícios e desafios



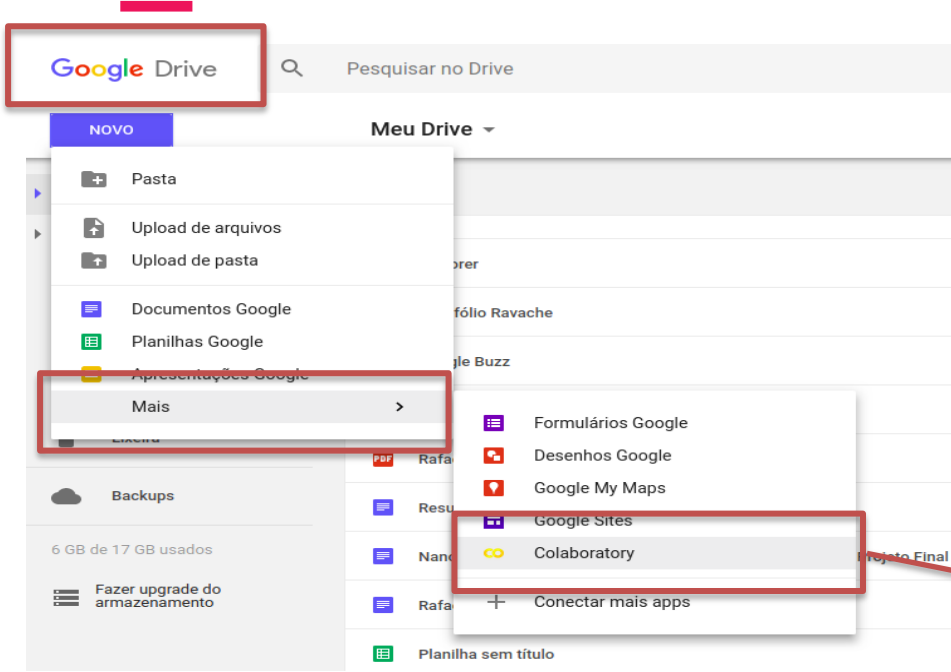
O **Google Colab** é uma plataforma cloud gratuita que te permite rodar código Python (2.7 e 3.6) em um formato muito parecido com um Jupyter Notebook de forma colaborativa. E agora disponibiliza o uso de GPU!



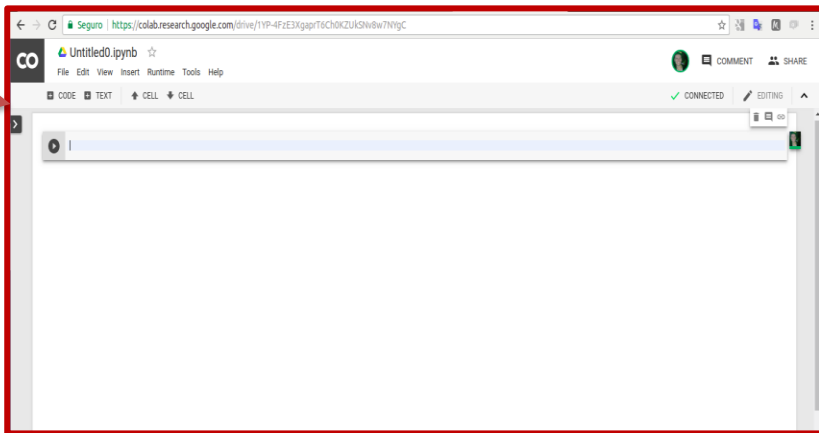
Python!

Obs.: Fique a vontade para desenvolver localmente também!

Como acessar o ambiente Cloud que vamos desenvolver?



The screenshot shows the Google Drive web interface. The 'Google Drive' logo is highlighted with a red box. Below it, the 'NOVO' (New) button is visible. The 'Mais' (More) option in the dropdown menu is also highlighted with a red box. A secondary dropdown menu is open, showing various Google apps, with 'Colaboratory' (represented by a yellow infinity icon) highlighted by a red box. A red arrow points from the 'Colaboratory' option to the Colab interface on the right.



The screenshot shows the Google Colaboratory (Colab) interface. The browser address bar displays the URL: <https://colab.research.google.com/drive/1YP-4FE3AgapTsCH0KZUK9N6w7NigC>. The interface includes a menu bar with options like File, Edit, View, Insert, Runtime, Tools, and Help. Below the menu bar, there are tabs for CODE, TEXT, and CELL. The main area is titled 'Untitled0.ipynb' and shows a code editor with a single cell containing a cursor. The interface is framed by a red border.

<https://colab.research.google.com/notebooks/intro.ipynb>

- NF (total final) = Nota do trabalho
- Construir um modelo de classificação de chamados

Criar um classificador de chamados aplicando técnicas de PLN

A **QuantumFinance** tem um canal de atendimento via chat e precisa classificar os assuntos dos atendimentos para melhorar as tratativas dos chamados dos clientes. O canal recebe textos abertos dos clientes relatando o problema e/ou dúvida e depois é direcionado para algum uma área especialista no assunto para uma melhor tratativa.

Crie um modelo classificador de assuntos aplicando técnicas de PLN, que consiga classificar através de um texto o assunto conforme disponível na base de dados [1] para treinamento e validação do modelo seu modelo.

O modelo precisa atingir um score na **métrica F1 Score superior a 75%**. Utilize o dataset [1] para treinar e testar o modelo, separe o dataset em duas amostras (75% para treinamento e 25% para teste com o `random_state` igual a 42).

Fique à vontade para testar e explorar as técnicas de pré-processamento, abordagens de NLP, algoritmos e bibliotecas, mas explique e justifique as suas decisões durante o desenvolvimento.

Composição da nota:

- **50%** - Demonstrações das aplicações das técnicas de PLN (regras, pré-processamentos, tratamentos, variedade de modelos aplicados, organização do pipeline, etc.)
- **50%** - Baseado na performance (score) obtida com a amostra de teste no pipeline do modelo campeão (validar com a Métrica F1 Score). **Separar o pipeline completo do modelo campeão conforme template.**

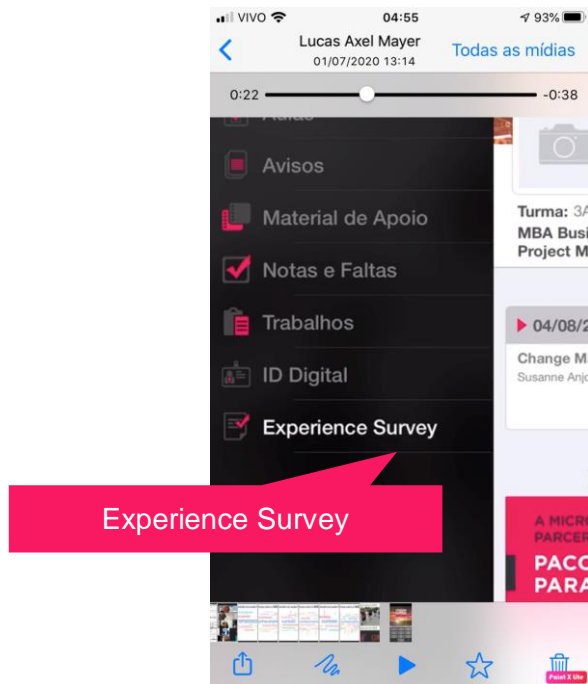
O trabalho poderá ser feito em grupo de até 4 pessoas (mesmo grupo do StartupOne).

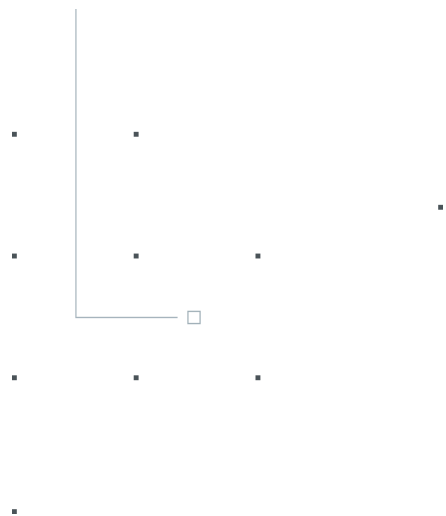
[1] = https://dados-ml-pln.s3.sa-east-1.amazonaws.com/tickets_reclamacoes_classificados.csv

Combinado é sempre avaliar minhas aulas, beleza?

Pelo aplicativo da FIAP ou pelo site

(Entrar no FIAP, e no menu clicar em Experience Survey)





Animados?

FIAP