人文社会ビジネス科学学術院 ビジネス科学研究群 2024年度 春C

テキストマイニングの実践

実践編のテーマ - ホテルのクチコミ分析 (by 楽天トラベル)

● ホテルのクチコミ数: 1,393万件 ※年間約60~80万



経年変化:

```
780万件 (2015)
```

- → 836万件 (2016)
- → 900万件 (2017)
- → 973万件 (2018)
- → 1,042万件 (2019)
- → 1,098万件 (2020)
- → 1,165万件 (2021)
- → 1,237万件 (2022)
- → 1,325万件 (2023)
- → 1,393万件 (今回)
 - ※ 2024/5/19現在

スケジュール

day 2

- ・ 講義 ー 自然言語処理の最新動向
- 講義 テキストマイニングの手順
- 演習 データ理解

day 3

- ・ 演習 / 演習環境の準備
- 演習 テキスト解析 (1)
- 演習 テキスト解析 (2)

day 4

- 講義&演習 データ分析 (使い方編)
 - 講義&講義一 データ分析 (実践編)

day 5

EXCEL

講義&講義― データ分析 (実践編)

SageMaker Studio Lab

● 近年、機械学習の教育・研究を目的とした研究用ツールが相次いで登場

演習で使用

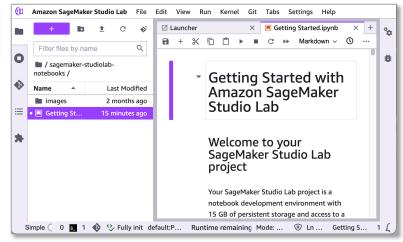


https://colab.researc h.google.com





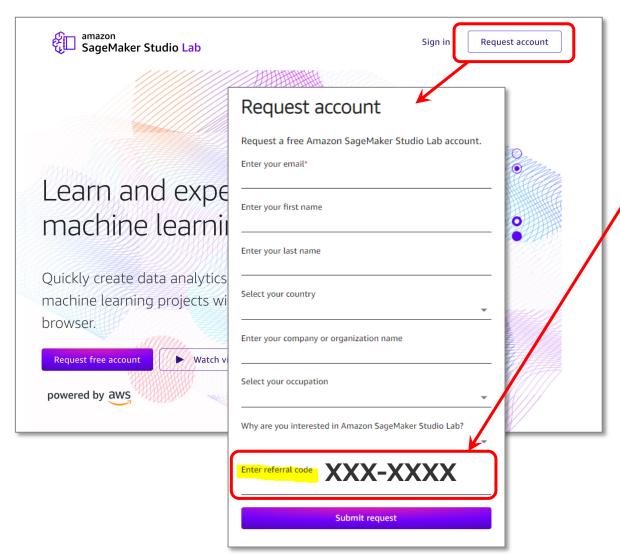
https://studiolab.sag emaker.aws/



	Clab(無償版)	Studio Lab
GPU	T4(16GB)	T4(16GB)
最長 実行時間	12時間	CPU:12時間 GPU:4時間
メモリ	12GB	15GB
ディスク	CPU:100GB GPU:78GB	15GB (永続化)
ターミナル	×	
ランタイムの 保存と再開	×	
費用	無償	無償
その他	Googleアカウン トが必要	AWSアカウント は不要 (クレカ 不要)

SageMaker Studio Lab のアカウント作成

● https://studiolab.sagemaker.aws/ にログインして、アカウントを作成してください



アカウント作成手順

- 1. <u>アカウント作成フォーム</u>からアカウントの申し込みを行う **注意**: **リファラルコード**をアカウント作成フォームに忘 れずに入力ください (**受講者限定**です)
- 2. 「Account request confirmed ...」のメールを受信し、メール内のリンクからアカウントを作成する→ リクエストの受付はすぐにメールが届きます
- 3. 「Verify your email ...」のメールを受信し、メール 内のリンクからメールアドレスを認証する → リファラルコードを利用している場合は2~3分以内 に結果が届きます
- 4. 「Your account is ready …」のメールを受信する

 → これで「Sign in」できます
 - ※ リファラルコードの有効期間: 2024/6/27 ~ 2024/7/12

R06年度 01KA438, 0ADM126 テキストマイニング