

CBE:5120 Data Science in Chemical and Engineering Systems

Fantastic Data and Where to Find [It]

Prof. Joe Gomes

October 2, 2020

Data Sources

- Google Datasets
- KAGGLE
- Misc.

Google Dataset Search

Dataset Search

Search for Datasets



Try [coronavirus covid-19](#) or [global temperatures](#).

[Learn more](#) about including your datasets in Dataset Search.

Kaggle

The screenshot displays the Kaggle homepage interface. On the left is a sidebar with navigation links: Home, Compete, Data, Notebooks, Discuss, Courses, Jobs, and More. The main content area features a search bar at the top. Below it, a dataset card for 'IMDB dataset (Sentiment analysis) in CSV format' is highlighted. This card shows it was created by Ziqi Yuan 10 months ago, has 16 code snippets, 0 discussions, and 9k views. It lists three CSV files: Test.csv, Train.csv, and Valid.csv, each with a brief description of the data. To the right of the dataset card, there's a user profile for Joe Gomes, who is an 'EXPERT' (Novice level) and joined 10 months ago. Below the profile is a 'Jobs' section listing several companies: Max Planck Institute for Astron..., Parenthetic, Washington, Distri..., and Reify Health, Boston, Massach... At the bottom right, there's a footer with '© 2020 Kaggle Inc' and links to 'Documentation', 'Our Team', and 'Terms'.

Navigation links: Home, Compete, Data, Notebooks, Discuss, Courses, Jobs, More

Search bar

Dataset: IMDB dataset (Sentiment analysis) in CSV format
IMDB Movie Review Dataset transform into CSV files

CSV | 65.9 MB | World Bank Dataset Terms of Use

16 | 0 | 9k

Data (3 files)

- Test.csv ()
Movie reviews are located in text field. Sentiment label are located in I...
- Train.csv ()
Movie reviews are located in text field. Sentiment label are located in I...
- Valid.csv ()
Movie reviews are located in text field. Sentiment label are located in I...

User: Joe Gomes
Joined 10 months ago
EXPERT
Novice

Jobs

- Max Planck Institute for Astron...
- Parenthetic, Washington, Distri...
- Reify Health, Boston, Massach...
- Reify Health, Boston, Massach...
- Reify Health, Boston, Massach...

© 2020 Kaggle Inc

Documentation
Our Team
Terms

Misc. Data Sources

- UCI ([ML](#) and [Cheminf.](#))
- [EPA Environmental Dataset Gateway](#)
- [Open-access medical image data](#)
- [NIST Materials Data Repository](#)
- [Citrine.io](#)

CBE:5120 Data Science in Chemical and Engineering Systems

Goodbye