KDnuggets

ChatGPT Cheat Sheet

ChatGPT is a large language conversational AI built by OpenAI. It was trained using Reinforcement Learning from Human Feedback, similar to InstructGPT. ChatGPT understands the prompt and provides detailed repose that can help you with research, coding, and various data science tasks.

Ideas

Dataset Suggestion
>>> I want to build a predictive model for image classifiers. Can you please suggest the five most relevant datasets for my use case?

Suggest Resources
>>> I would like to learn about deep learning. Please suggest 3 best specific resources.

AB Testing
>>> [case-study]. Please design an A/B test for this purpose. Please include the concrete steps on which statistical test I should run.

Career Coaching
>>> I am looking for a role as a data engineer. My background is management. What should I do in 6 months to get a job?

Coding

Unit Test
>>> Write a unit test for train function. The test cases are: x should not be null value and y should be a numerical value.

Code Explanation
>>> Can you explain what the code is doing? `[code-snippet]`

Optimize Code
>>> Can you improve the time complexity of the code? `[code-snippet]`

SQL

SQL Formatting
>>> Format the following SQL code and convert all reserved keywords to uppercase. `[code-snippet]`

Translate Between DBMS
>>> What is the equivalent of PostgreSQL's DATE_TRUNC for MySQL?

Calculate Average
>>> Write the SQL code that works for PostgreSQL 14. I have a table with two columns [date, temp]. I would like to calculate an average temp.

Calculate Runway
>>> Write SQL to calculate my runway.

Spreadsheets

Spreadsheets Formula
>>> Create a spreadsheet formula to calculate the sum of cells B1 to B20?

Dummy Data
>>> Generate the dummy data for me to use as placeholders in my spreadsheet.

Tips
>>> Give me some tips on how to improve the efficiency of my spreadsheet?

Data Analysis

Generate Data
>>> Generate a fake data with 100 rows and 4 columns: [id,name,grade,subject]

Data Cleaning
>>> I have a text classification dataset. Write Python code for data cleaning.

Data Exploration
>>> I have a dataset of 100 rows and four columns:[id, name, grade, subject]. Write R code for data visualization and exploration.

Data Visualization
>>> I have a dataset with 100 rows columns [id, name, grade, subject]. Create a matplotlib bar chart of subject vs. grade.

Machine Learning

Train Regression Model
>>> You are a data scientist, write Python code for me. I have a dataset with columns [model,hp,speed]. Please build a machine learning model that predicts speed.

Hyperparameter Tuning
>>> I have a logistic regression model, write Python code to tune hyperparameters.

Imbalance Data
>>> I have an imbalanced dataset with target column species. In python, how do I oversample and/or undersample my data?

Explain the Model
>>> I have trained a LightGBM model. Write a Python code to explain the output using a series of plots with Shap.

Research

Explain the Concept
>>> Explain t-test to an undergraduate as a data science instructor.

Stakeholders
>>> Tips on explaining data science reports to a business stakeholder.

Summarize the paper
>>> Please summarize the paper "Adding Conditional Control to Text-to-Image Diffusion Models" in simple terms in one paragraph.

Writing Blog
>>> Write an outline for a blog "Python lists".

Research History
>>> Can you research the history of the graph neural network?

Subscribe to KDnuggets News