

starcoder2




Model Overview

- Possibilities
- Features
- Architecture
- Uses[°]
- Application
- Purpose



Introduction

What is StarCoder2 ?



{ In the dynamic world of technology, where innovation is the key to success, a triumvirate of tech giants — ServiceNow, Hugging Face, and NVIDIA — have joined forces to create a groundbreaking Large Language Model (LLM) for code generation, known as StarCoder2. This model is the fruit of a collaborative effort with the BigCode Community, a consortium managed by ServiceNow, a pioneer in digital workflow solutions, and Hugging Face, the most widely used open-source platform. The development of StarCoder2 was driven by a shared vision to revolutionize the field of code generation and set new benchmarks in performance, transparency, and cost-effectiveness. }

{ StarCoder2 is not just another LLM; it's a family of open-access models for code generation that are designed to redefine the standards of performance. It comes in three distinct sizes, each tailored to specific needs: a 3-billion-parameter model developed by ServiceNow, a 7-billion-parameter model crafted by Hugging Face, and a 15-billion-parameter model engineered by NVIDIA. Each variant is a testament to the collaborative spirit and technical prowess of the teams involved }

StarCoder2 is a family of code generation models (3B, 7B, and 15B), trained on 600+ programming languages from The Stack v2 and some natural language text such as Wikipedia, Arxiv, and GitHub issues. The models use Grouped Query Attention, a context window of 16,384 tokens, with sliding window attention of 4,096 tokens. The 3B & 7B models were trained on 3+ trillion tokens, while the 15B was trained on 4+ trillion tokens. For more details check out the

Still in development: StarCoder2 is a new language model and is still in development.

May not be accurate in all cases: While StarCoder2 is highly accurate, it may not be accurate in all cases, especially with complex code. The StarCoder2 model was developed in 2022. It was first announced at the NeurIPS 2022 conference, and has been available for trial use since then.



Purpose

Its purpose is to facilitate programming processes, make them less complex, and accelerate the solution and development of software problems



Possibilities



Code Perfection: StarCoder2 can predict the next code in a given context, helping developers write code faster and more efficiently.

Code translation

StarCoder2 can convert code from one programming language to another, making it easier for developers to work with different languages.

Bug Fixing:

StarCoder2 can help developers find and fix bugs in their code.

Code Generation



Features

Accurate: StarCoder2 is trained on a huge dataset of code, making it extremely accurate in understanding and generating code.

Fast: StarCoder2 can process code very quickly, making it an effective tool for developers.

Easy to use: StarCoder2's interface is designed to be easy to use, even for novice developers

Starcoder-2

Download ▾



Demo

API

Documentation

Description

StarCoder is a LLM specializing in code completion.

Publisher

BigCode

Latest Version

-

Modified

February 27, 2024

Size

0 B

AI Foundation Model

By using this demo, you acknowledge that you have read and agreed to the [terms & conditions](#).

M

X_train, y_train, X_test, y_test = train_test_split(X, y, test_size=0.1) #Train a logistic regression model, predict the labels on the test set and compute the accuracy score

Temperature : 0.2 Top P : 0.7 Max Tokens : 1024



S

```
#+
#Import the necessary modules
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import confusion_matrix, classification_report

#Create training and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.4, random_state=42)

#Create the classifier: logreg
logreg = LogisticRegression()

#Fit the classifier to the training data
```

What is Earth's relationship to the Sun?

Submit

Enter a prompt to receive an AI-generated text.

Show Parameters

JAVA



C++

CONNECTED

Uses

Software Development: StarCoder2 can be used to help developers at all stages of the software development process, from writing code to testing it.

Education: StarCoder2 can be used to teach new developers how to write code.

Application



References

Medium

Huggingface

Github

Nvidia

Big code



summary

StarCoder2 is a promising tool for developers and researchers working with code. Its code completion, translation, error detection, and generation capabilities make it a valuable tool for increasing productivity and improving code quality. However, it is important to remember that it is still under development, and may not be perfect in all cases.

As of April 2024, the official StarCoder2 website may not yet be fully available to the public. However, you can stay tuned for announcements from BigCode or Google AI regarding its wider release. Alternatives: In the meantime, many other code completion and generation tools are available, such as GitHub Copilot, Kite, and DeepCode.

Thank you!

 Presented by : Safa Nasser-Layan Alharbi