CS-C1000 – Introduction to Artificial Intelligence

Computer Assignment B

William Wilkinson

March 23, 2021

Computer Exercise B: Fake Text Generation

Artificial intelligence (AI)

New AI fake text generator may be too dangerous to release, say creators

The Elon Musk-backed nonprofit company OpenAI declines to release research publicly for fear of misuse

Alex Hern

@alexhern

Thu 14 Feb 2019 17.00 GMT

21,971 572

▲ The AI wrote a new passage of fiction set in China after being fed the opening line of Nineteen Eighty-Four by George Orwell (pictured). Photograph: Mondadori/Getty Images

The creators of a revolutionary AI system that can write news stories and works of fiction – dubbed "deepfakes for text" – have taken the unusual step of not releasing their research publicly, for fear of potential misuse.

OpenAI, an nonprofit research company backed by Elon Musk, Reid Hoffman, Sam Altman, and others, says its new AI model, called GPT2 is so good and the risk of malicious use so high that it is breaking from its normal practice of releasing the full research to the public in order to allow more

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

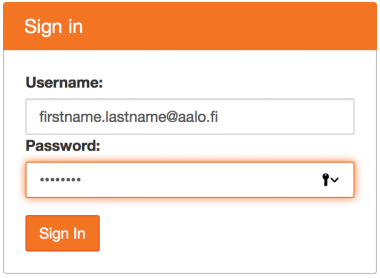
Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns.

Log into JupyterHub (a quick reminder)

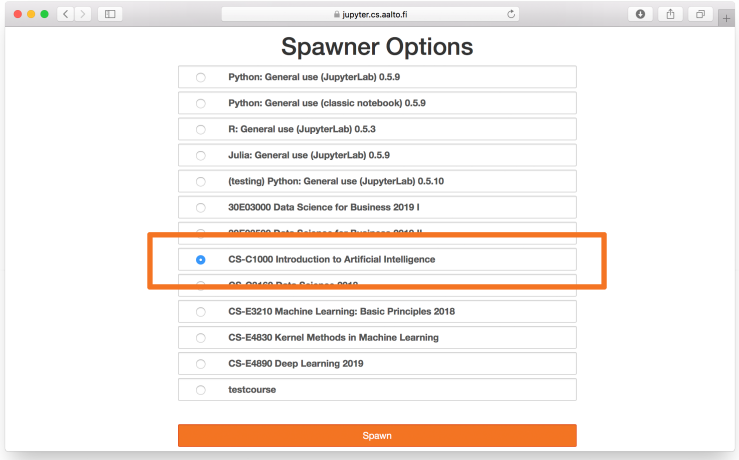
Log in to:

`https://jupyter.cs.aalto.fi`

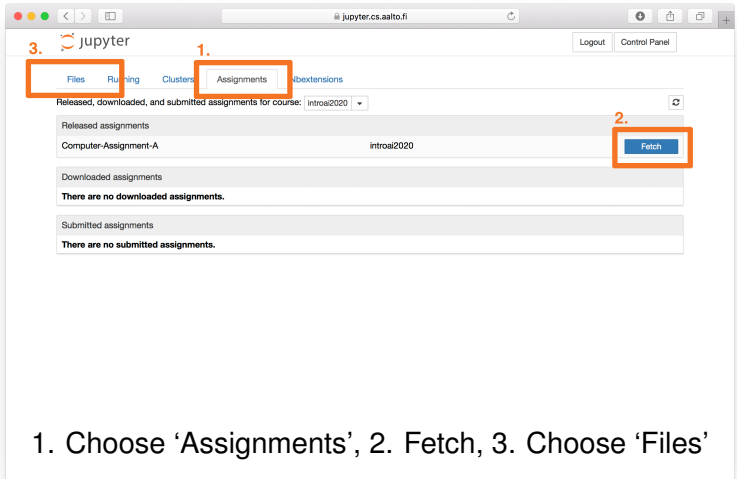


Using your Aalto account.

Choose the course

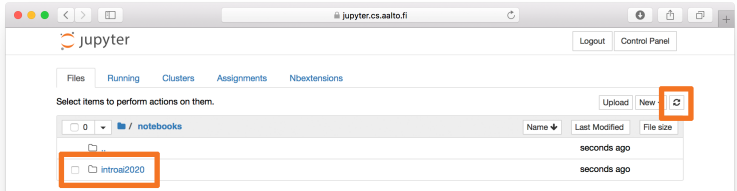


Fetch the exercise (B this time, not A!)



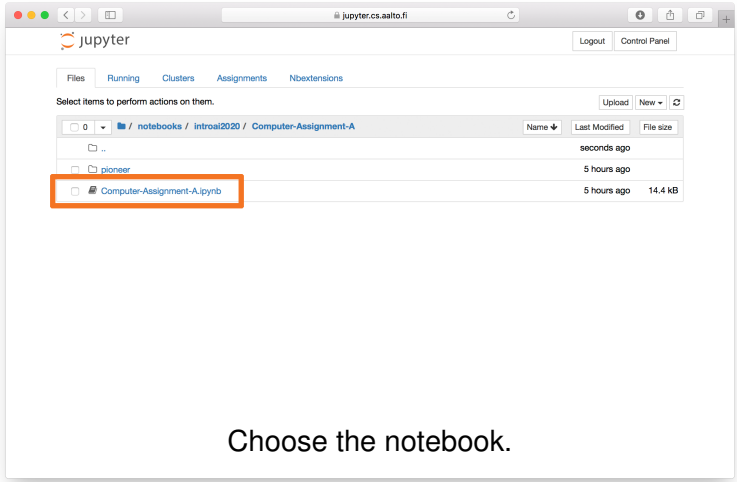
1. Choose ‘Assignments’, 2. Fetch, 3. Choose ‘Files’

Find the notebook



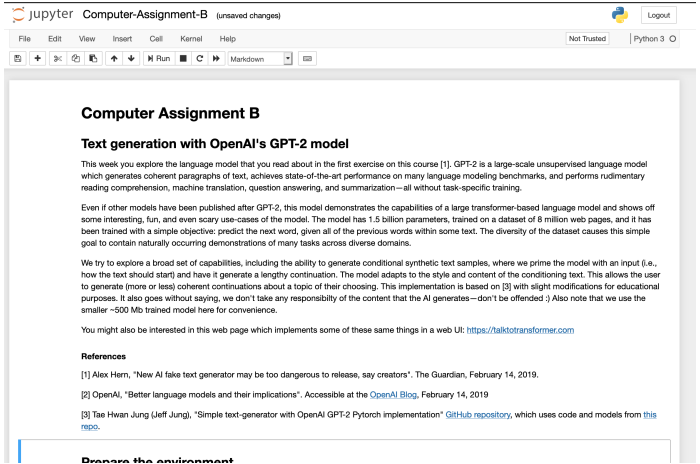
Refresh the view if the course does not show.

Fire it up (again, B this time)



Choose the notebook.

Ready to run (a cell at a time)



Tasks (read the notebook for exact instructions)

- Generate some unconditional fake text.
- Auto-complete some text.
- Generate multiple novel text samples.
- Generate samples of varying length.
- Modify the model parameters.

Tasks (read the notebook for exact instructions)

- Generate some unconditional fake text.
- Auto-complete some text.
- Generate multiple novel text samples.
- Generate samples of varying length.
- Modify the model parameters.

Tasks (read the notebook for exact instructions)

- Generate some unconditional fake text.
- Auto-complete some text.
- Generate multiple novel text samples.
- Generate samples of varying length.
- Modify the model parameters.

Tasks (read the notebook for exact instructions)

- Generate some unconditional fake text.
- Auto-complete some text.
- Generate multiple novel text samples.
- Generate samples of varying length.
- Modify the model parameters.

Tasks (read the notebook for exact instructions)

- Generate some unconditional fake text.
- Auto-complete some text.
- Generate multiple novel text samples.
- Generate samples of varying length.
- Modify the model parameters.

Tasks (read the notebook for exact instructions)

- Generate some unconditional fake text.
- Auto-complete some text.
- Generate multiple novel text samples.
- Generate samples of varying length.
- Modify the model parameters.

How to get points?

- To get the points for this week's exercise, answer the questions in MyCourses related to this Computer Exercise B.