# Readme

# This is a Image captioning model I have made for week 4 of CSOC, as my final project. This model generates a story on the by extracting the features of an input Image[¶](https://kkb-production.jupyter-proxy.kaggle.net/static/dist/jupyterlab/jupyterlab-index-560be39f31dae074b6dc.html?session=eyJhbGciOiJub25lIiwidHlwIjoiSldUIn0.#This-model-generates-a-story-on-the-by-extracting-the-features-of-an-input-Image)

# So basically, it consists of 3 submodels in it:

# 1. Feature extraction model by extracting a layer from a image classification model (Computer Vision)

# 2. LSTM model, trained on a dataset with predefined captions (Natural Language Processing)

# 3. A Top-p (nucleus) sampling model using transformers, which generate a story from the captions generated (Natural Language Processing)

# I have done this in 2 files. In the first model, “features.ipynb”, I have trained an Image classifier from scratch using RESNET architecture, and then downloaded the extracted features for all images in the dataset in "custom\_resnet\_features.pkl", which is then passed into the next model, in the next jupyter notebook, “nlp.ipynb”, where other 2 models are trained.