



Capstone Project – Neural Networks

Task

[Visit our website](#)

Introduction

Great job! You have reached another capstone project in this bootcamp. This project provides you with an opportunity to apply sequential modelling to a natural language processing (NLP) problem. This will equip you with some of the necessary skills for real-world NLP challenges. You will have to showcase your data preprocessing, model training, and knowledge synthesis skills.

Developer Portfolio

Developers who have the edge are those who find ways to apply their newfound skills from the get-go. As you may know, a **developer portfolio** (a collection of online creations that you have made) allows you to demonstrate your skills rather than just telling people about them. It's a way of bringing your CV to life and introducing yourself to the world. As you learn more skills and put these into practice, each project that you complete will become more efficient and eye-catching.

These capstone projects give you the means to create projects for your very own developer portfolio, allowing you to walk away from this course not only with a certificate but, more importantly, with a head start on your tech career!

Task At Hand

In this task, you will be creating a program that helps businesses generate slogans based on their industry and predict the industry of a business based on a given slogan. A Jupyter Notebook template has been provided to guide you through this project. Copy the template program provided, `capstone_template.ipynb`, and rename it `capstone_project.ipynb`. This template has been provided to make this task a little easier for you. Your job is to open the template, follow all the instructions, and complete the steps to build both the slogan generator and classifier. Remember to save your work as you go along.

Instructions

A key focus of this project will be ensuring that your code is correct, well-formatted, and readable and that it adheres to the [PEP 8 style guide](#). In this regard, make sure that you do the following (and double-check before submitting your work to avoid losing marks unnecessarily!):

1. Identify and remove all syntax, runtime, and logical errors from your code.
2. Make sure that your code is readable. To ensure this, add comments to your code, use descriptive variable names, and make good use of whitespace and indentation.
3. Make sure that your code is modular. Create functions to perform specific units of work.
4. Make sure that your code is as efficient as possible. How you choose to write code to create the solution to the specified problem is up to you. However, make sure that you write your code as efficiently as possible.
5. Make sure that all outputs that your program provides to a user are easy to read and understand. Label all data that you output.



Practical task

This project involves working with a dataset containing various business slogans and their corresponding industries. The **task template** provided will walk you through each of these steps. Work through them in order, and do not edit the provided code.

- Make a copy of the task template and rename it **neural_network_task.ipynb**.
- We recommend using Google Colab for this task but you can also do it on your local machine if you choose to do so. Note that **spacy** is already installed in Google Colab.

Your task is divided into four main parts:

1. Data Preprocessing:

- Load the slogan dataset.
- Extract relevant columns and handle missing values.

- Tokenise the slogans and business names in the dataset (code provided).

2. Slogan Generator:

- Build an LSTM model that can generate slogans based on an input industry.
- The model will take the industry category (encoded as a numerical value) as input and produce a generated slogan as output.
- Train your model on the provided slogan dataset and test it by generating slogans for different industries.

3. Slogan Classifier:

- Create a classification model that predicts the industry of a business based on its slogan.
- Use the tokenised and encoded slogans as inputs, and predict the corresponding industry category.
- Split your dataset into training and testing sets and evaluate how well the model performs.

4. Combining Both Models:

- Use the generator to create new slogans for a given industry.
- Pass the generated slogans through the classifier to see if it can correctly predict the industry.
- Compare the results and comment on any differences you notice between the generated slogans and the classifier's predictions.

Important: Be sure to upload all files required for the task submission inside your task folder and then click "Request review" on your dashboard.



Share your thoughts

Please take some time to complete this short feedback [form](#) to help us ensure we provide you with the best possible learning experience.