Section 1: Project Definition

Project Overview: state the high-level overview of the project, including the background information such as problem domain, project origin, and related data sets or input data.

The goal of this project is to classify images of dogs according to their breed using deep learning and in particular convolutional neural networks.

The input data consists of a dataset of dog images which is provided. This dataset is already divided in train, validation and test images. There are 133 total dog categories and 8351 total dog images. There are 6680 training dog images, 835 validation dog images, and 836 test dog images.

Problem Statement: define the problem to be solved.

I will model and train several CNN algorithms that will used this dataset to learn how to predict a dog breed when a new image is provided. I will also use these algorithms to make prediction for new data. Convolutional Neural Networks consist of several layers with small neuron collections, each of them perceiving small parts of an image. The results from all the collections in a layer partially overlap in a way to create the entire image representation. The layer below then repeats this process on the new image representation, allowing the system to learn about the image composition. The Convolutional Neural Network (CNN or ConvNet) is a subtype of Neural Networks that is mainly used for applications in image and speech recognition. Its built-in convolutional layer reduces the high dimensionality of images without losing its information.

Metrics: define the metrics to measure the results and justifications to use the metrics. For example, if you use time-series data sets, what metrics will be appropriate to measure the results.

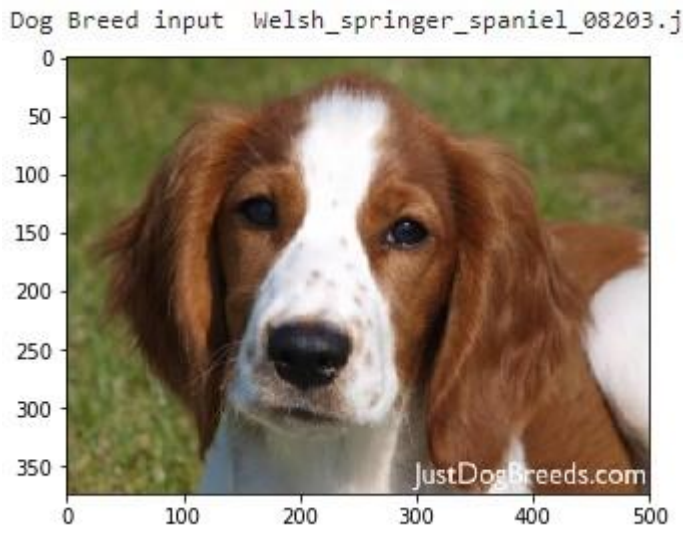
The metric I will use to test how good the algorithm is will be the accuracy. I already have train, validation and test data. Then, I will use the train data to train the algorithms. After, I will use the validation data to decide between several model options which is the best one. Finally, I will use the test data to see how good the algorithm is predicting results for unobserved data.

Section 2: Analysis

Data Exploration: describe the data sets, including the features, data distributions, and descriptive statistics. Identify any abnormalities or specific characteristics inherent in the data sets.

The dataset consists in dog images, as well as the label for each image indicating the dog breed.

Data Visualization: build data visualization based on the data exploration in the previous step.



Section 3: Methodology

Data Preprocessing: describe the steps taken to preprocess the data and address any abnormalities in the data sets. If data preprocessing is not needed, please explain why.

Data preprocessing is not needed because the input data was already provided in an appropriate format to be used.

Implementation: discuss the process using the models, algorithms, and techniques applied to solve the problem. Any complications during the implementation should be mentioned.

I have implemented three CNN algorithms: I have modelled one algorithm from scratch, and two algorithms using transfer learning: VGG16 and Resnet50.

Refinement: describe the process to refine the algorithms and techniques, such as using cross-validation or changing the parameter settings.

I have trained each algorithm during several epochs. I have checked which one was getting higher accuracy using cross-validation with the validation data, and I have selected the ones that got the highest.

Section 4: Results

Model Evaluation and Validation: discuss the models and parameters used in the methodology. If no model is used, students can discuss the methodology using data visualizations and other means.

- Algorithm from scratch: I have added three sequential combinations of a 2D convolution layer and a 2D max pooling layer. The convolution layer will apply a filter to an input to create a feature map that summarizes the presence of detected features in the input. The max pooling layer will select the maximum element from the region of the feature map covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previous feature map. Then have added a 2D global average pooling, that will compute the average value of all values across the entire matrix for each of the input channels. And the final layer, or dense layer, which is a layer that is deeply connected with its preceding layer which means the neurons of the layer are connected to every neuron of its preceding layer.
- Algorithm VGG16: The CNN architecture consists on a convolutional neural network that is 16 layers deep, and it has been trained on more than a million images from the ImageNet database. Then after that part, the 2D global average pooling layer is added, which is used for applying global average pooling operation for spatial data. And the final layer, or dense layer, applies the softmax activation function for multi-class classification problems, and for each image predict the output class with a higher value. And it is trained using our dogs dataset.
- Algorithm Resnet50: The CNN architecture is based on the VGG-16 convolutional neural network. Then after that part, the 2D global average pooling layer is added, which is used for applying global average pooling operation for spatial data. And the final layer, or dense layer, applies the softmax activation function for multi-class classification problems, and for each image predict the output class with a higher value. And it is trained using our dogs dataset.

Justification: discuss the final results in detail and explain why some models, parameters, or techniques perform better over others. Show and compare the results in tabular forms or charts.

Algorithm	**Accuracy**
From scratch	1%
VGG16	40%
Resnet50	80%

The algorithms using transfer learning perform much better than the one build from scratch. Because they have already been pre trained to classify images. Then the training needed consists on learning about the specifics images in each use case, in these particular one focusing only in dog images to predict the dog breed.

Section 5: Conclusion

Reflection: summarize the end-to-end problem solution and discuss one or two particular aspects that you find interesting or difficult to implement.

After comparing the three algorithms, I choose to use the Resnet50 because is the one that gets the higher accuracy by far.

Modelling and algorithm from scratch has been the most difficult part for me, because I did not know how many layers and which kind of layers I needed to use for this use case, and I was getting really low accuracy, That is why at the end I decided to model the suggested architecture. Still as we can see the accuracy is incredibly low.

I find really interesting that using a pretrained algorithm with images not related to our problem could give so good results for ours.

Improvement: provide suggestions for the next research to improve the experiment.

Some possible improvements that can be applied:

- Data augmentation, so we can have similar images to the ones provided but in different positions; and we will have a bigger input data set.
- Editing images and changing color or size, or displaying only parts of the original image; then the algorithm is trained when images are not displaying the full dog.
- Modifying the algorithm with more hidden layers, training with more epochs; maybe the algorithm needs more training to perform better.