NEURAL NETWORK MODELS FOR OBJECT RECOGNITION

University: University of Essex Online

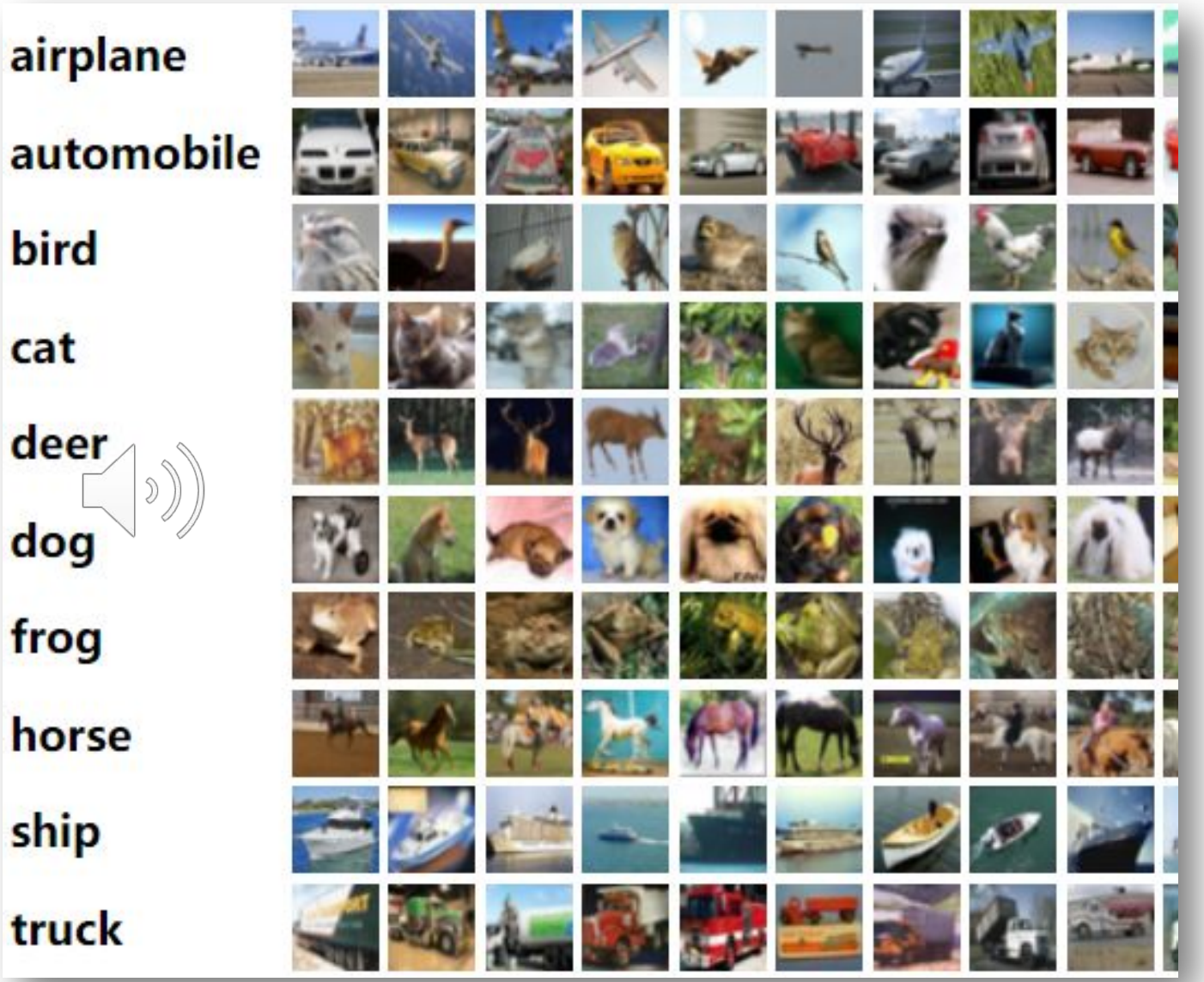
Module: Machine Learning

Unit 11: Neural Network models for object recognition

Team members: Samuel Adeniyi, Rachel Mead and Konstantinos Kyriacou

THE DATASET-CIFAR10

- **The dataset** consists of 60,000 images, each belonging to one of the following ten categories: ‘airplane’, ’automobile’, ’bird’, ‘cat’, ’deer’, ‘dog’, ‘frog’, ‘horse’, ‘ship’ and ‘truck’. Thus, there are 6,000 images per category (Canadian Institute for Advanced Research, 2009)



OBJECTIVES

OBJECTIVES

- Build a Convolutional Neural Network using the CIFAR-10 dataset
- Evaluate the performance of the convolutional neural network
- Build a model with sufficiently high performance to enable each image to be classified correctly, according to their respective category

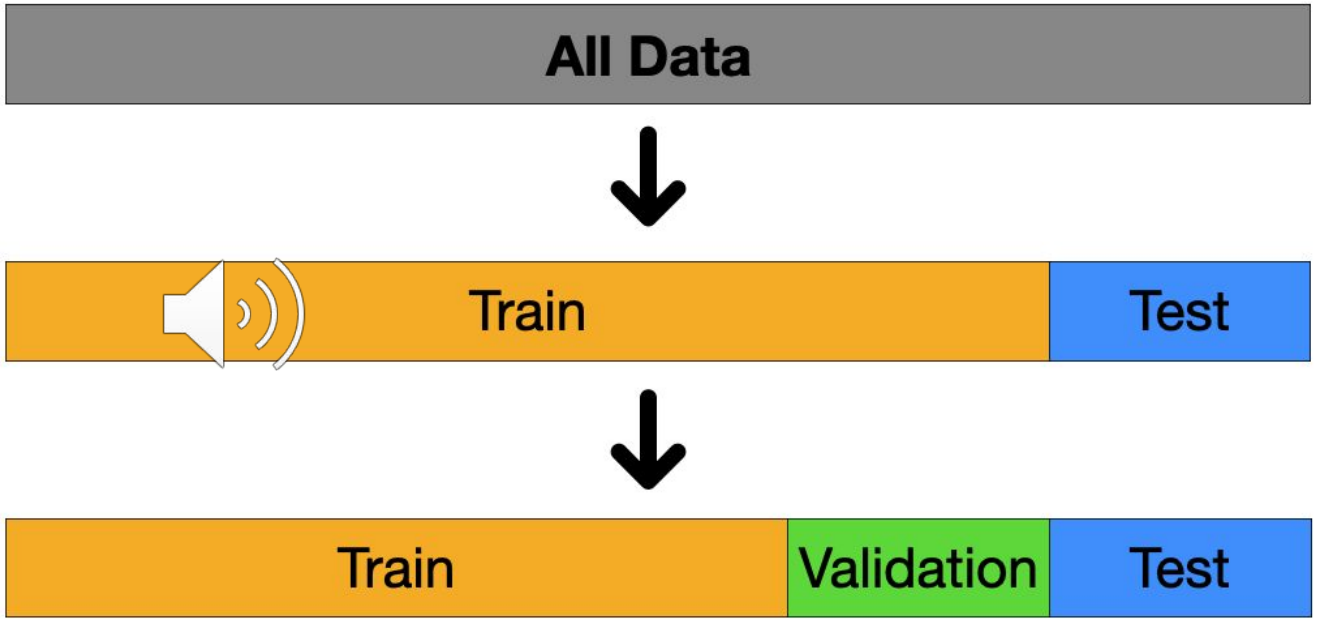
TRAINING AND VALIDATION SET

TRAINING SET:

- Train the model
- Provide the best fit of the model
- Improve the prediction rate when predicting the test samples

VALIDATION SET:

- Select the initial 10,000 records of the training set to be utilized for our validation procedure



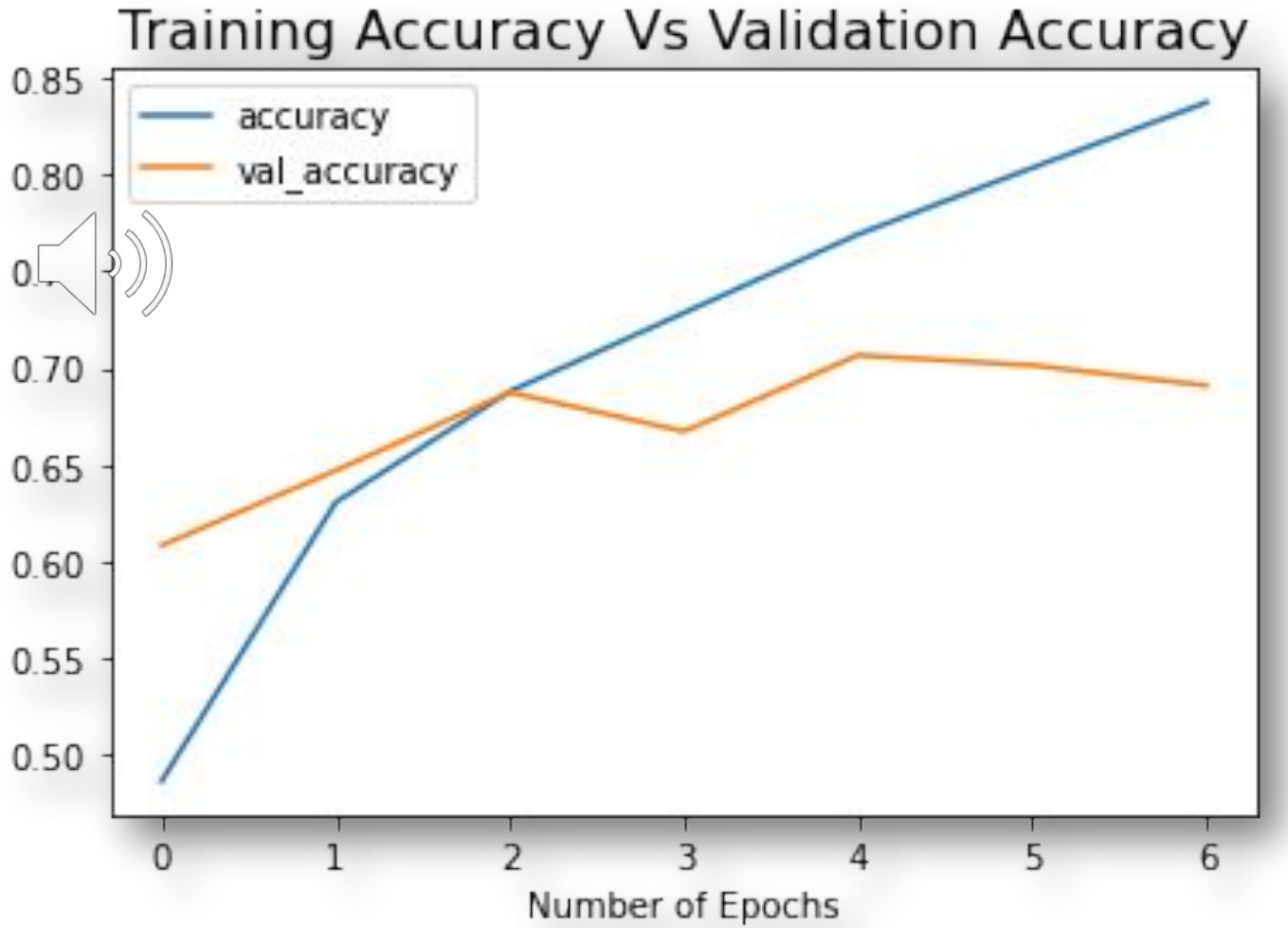
TRAINING MEASURES VS VALIDATION MEASURES

TRAINING MEASURES

- **The training loss** demonstrates the measure of deviation of the model from the training data
- **The validation loss** demonstrates the measure of deviation of the model when fitting new data (Yathish, 2022)

VALIDATION MEASURES

- **The training accuracy** demonstrates the measure of accuracy of the model on the training data
- **The validation accuracy** demonstrates the measure of accuracy when modelling new data



VALIDATION SET

- **Review the performance of the model on new data** based on different combinations of hyperparameters, thus to proceed to the best choice of modelling for predicting the test data

- **Prevent the model from overfitting,** meaning that the model performed well in classifying the training data but, on the other hand, was not able to make accurate predictions with data that was as yet unseen by the model

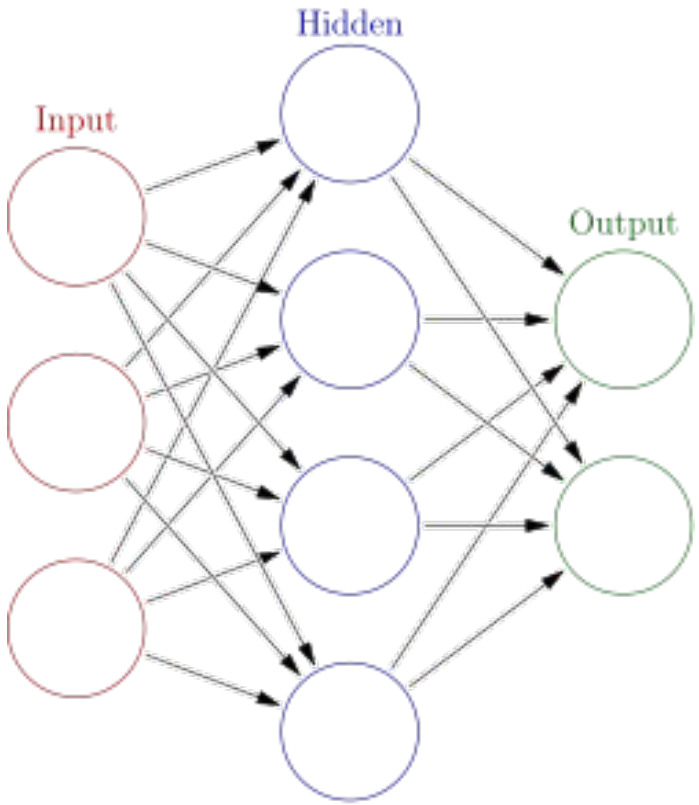
THE STRUCTURE OF ANN

ADVANTAGES

- Gives the ability of the Artificial Neural Network to extract the feature of the input more efficiently
- Improves the performance of the model
- Allows for parallel processing and for more efficient co utation
- Has the ability to adjust and provides great learning potential

DISADVATAGES

- It is computationally costly
- The potential for overfitting is greater in comparison to other modelling structures



ACTIVATION FUNCTIONS

ReLU

Advantages:

- The output of the function is not linear which increases the accuracy of the model
- Train faster the model
- Less computationally expensive

Disadvantages:

- Some neurons may stop responding to variations in error or input along the process of modelling (DanB, 2018)

SoftMax

Advantages:

- Contribute in determining the final decision as regards the classification of the image
- Weight the max logit significantly greater than the rest logits (Wood, N.D.)

CATEGORICAL CROSS ENTROPY LOSS FUNCTION

Advantages

- Easy to implement and optimize
- The graphical representation of the function consists of a smooth and convex shape (Koech, 2020; Kumar 2020)
- It is not affected when scaling or shifting of the predicted probabilities if the probabilities values still range between 0 and 1 (LinkedIn, 2023)

Disadvantages

- Sensitive to outliers or imbalanced data
- The accuracy of the model performance is not indicated directly (Brownlee, 2019)

NUMBER OF EPOCHS UTILISED

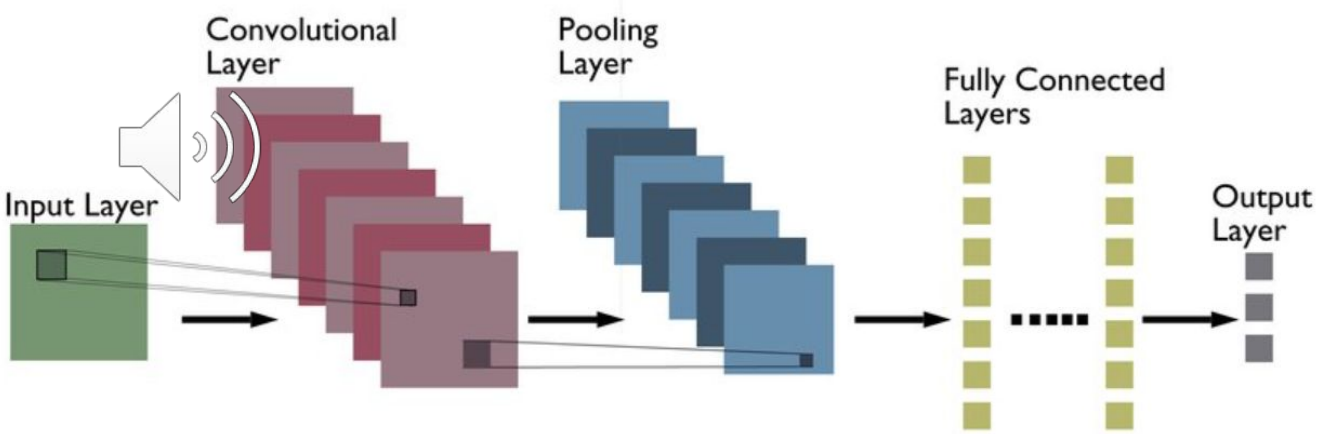
- **25 epochs** were selected to train our model and thus improve its performance for every subsequent iteration.
- **<u>Constraint</u>:** When the model performance reaches a point where the outcome of the cross-entropy loss function for the validation data remains the same for the two subsequent cycles then we conclude that our model has reached at a satisfactory level of performance and thus the training should stop at that point. Consequently, the model training process performs only the initial 7 iterations until it reaches that point, thus the rest cycles are omitted.

```
history = model.fit(X_train,y_categorical_train,epochs=25,validation_data=(x_validation,y_validation_categorical),callbacks=

Epoch 1/25
1250/1250 [==============================] - 144s 113ms/step - loss: 1.4288 - accuracy: 0.4866 - val_loss: 1.0968 - val_accura
cy: 0.6082
Epoch 2/25
1250/1250 [==============================] - 124s 99ms/step - loss: 1.0548 - accuracy: 0.6306 - val_loss: 1.0031 - val_accurac
y: 0.6469
Epoch 3/25
1250/1250 [==============================] - 129s 103ms/step - loss: 0.8960 - accuracy: 0.6877 - val_loss: 0.9103 - val_accura
cy: 0.6873
Epoch 4/25
1250/1250 [==============================] - 128s 102ms/step - loss: 0.7768 - accuracy: 0.7283 - val_loss: 0.9569 - val_accura
cy: 0.6671
Epoch 5/25
1250/1250 [==============================] - 129s 104ms/step - loss: 0.6643 - accuracy: 0.7689 - val_loss: 0.8732 - val_accura
cy: 0.7064
Epoch 6/25
1250/1250 [==============================] - 122s 97ms/step - loss: 0.5617 - accuracy: 0.8032 - val_loss: 0.8912 - val_accurac
y: 0.7012
Epoch 7/25
1250/1250 [==============================] - 110s 88ms/step - loss: 0.4683 - accuracy: 0.8370 - val_loss: 1.0096 - val_accurac
y: 0.6908
```

NEURAL NETWORK DESIGN

- During the initial phase of modelling, double sets of layers are constructed, each one consisting of a convolutional layer and a pooling layer
- The convolutional layer is utilized to extract several features from the input image. Following that, the output becomes the input of the following pooling layer whose primary aim is to reduce the size of the convolved feature map and thus reduce the computational costs (Kumar, 2021)



NEURAL NETWORK DESIGN

Convolutional Layer

- The input shape consists of three dimensions which are equal to 32 x 32 x 3
- The length and width of the kernel are both chosen to be equal to 3
- The stride hyperparameter is chosen to be equal to 1
- The number of filters hyperparameter is chosen to be equal to 64 (Pradeep, 2017)

Max-Pooling Layer

- The length and width of the kernel are both chosen to be equal to 2

```
# FIRST SET OF LAYERS
# Add a convolutional layer
model.add(Conv2D(filters=64,strides=1,kernel_size=(3,3),activation='relu',input_shape=(32,32,3)))
# Add a pooling layer
model.add(MaxPooling2D(pool_size=(2,2)))

#SECOND SET OF LAYERS
#Add another convolutional layer
model.add(Conv2D(filters=64,strides=1,kernel_size=(3,3),activation='relu'))
# Add another pooling layer
model.add(MaxPooling2D(pool_size=(2,2)))
```

KNOWLEDGE GAINED

To conclude, the project has given us the opportunity to obtain extended knowledge as regards the implementation of Convolutional Neural Networks modeling for image classification purposes. However, this modelling technique has a wider range of applications including decoding facial recognition, analyzing documents or even understanding climate change (Kim et al, 2022).

There is a lot more to learn about this modelling technique and thus this project was just the beginning.


Thank you
For your
Attention

REFERENCES

- Canadian Institute for Advanced Research (2009) The CIFAR-10 dataset. Available from: https://www.cs.toronto.edu/~kriz/cifar.html [Accessed 16 July 2023].
- Yathish, V. (2022). Available from: https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9 [Accessed 16 July 2023].
- DanB (2018) Rectified Linear Units (ReLU) in Deep Learning. Available from: https://www.kaggle.com/code/dansbecker/rectified-linear-units-relu-in-deep-learning/notebook [Accessed 16 July 2023].
- Wood, T. (publication date unknown). What is the Softmax Function? Available from: https://deepai.org/machine-learning-glossary-and-terms/softmax-layer [Accessed 16 July 2023].
- Koech, K. E. (2020) Cross-Entropy Loss Function. Available from: https://towardsdatascience.com/cross-entropy-loss-function-f38c4ec8643e [Accessed 16 July 2023].
- Kumar, A. (2020) Keras – Categorical Cross Entropy Loss Function. Available from: https://vitalflux.com/keras-categorical-cross-entropy-loss-function/#:~:text=categorical_crossentropy%3A%20Used%20as%20a%20loss,into%20categorical%20encoding%20using%20keras [Accessed 16 July 2023].

REFERENCES

- LinkedIn (2023) What are the advantages and disadvantages of using cross-entropy loss for classification tasks? Available from: https://www.linkedin.com/advice/0/what-advantages-disadvantages-using-cross-entropy [Accessed 16 July 2023].
- Brownlee, J. (2019) A Gentle Introduction to Cross-Entropy for Machine Learning. Available from: https://machinelearningmastery.com/cross-entropy-for-machine-learning/ [Accessed 16 July 2023].
- Kumar, P. (2021) Max Pooling, Why use it and its advantages. Available from: https://medium.com/geekculture/max-pooling-why-use-it-and-its-advantages-5807a0190459 [Accessed 16 July 2023].
- Pradeep, S. (2017) Why would one use larger strides in convolutional NNs, as opposed to smaller strides? Available from: https://www.quora.com/Why-would-one-use-larger-strides-in-convolutional-NNs-as-opposed-to-smaller-strides [Accessed 16 July 2023].
- Brownlee, J. (2019) A Gentle Introduction to Pooling Layers for Convolutional Neural Networks. Available from: https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/ [Accessed 16 July 2023].
- Kim, J., Lee, M., Han, H., Kim, D., Bae, Y. and Kim, H.S. (2022). Case study: Development of the CNN model considering teleconnection for spatial downscaling of precipitation in a climate change scenario. Sustainability, 14(8), p.4719.