

1. Replicate the [notebook](#) on the next character prediction and use it for generation of text. Use one of the datasets specified below for training. Refer to Andrej Karpathy's blog post on [Effectiveness of RNNs](#). Visualise the embeddings using t-SNE if using more than 2 dimensions or using a scatter plot if using 2 dimensions. Write a streamlit application which asks users for an input text and it then predicts next k characters [5 marks]

Datasets (first few based on Effectiveness of RNN blog post from Karpathy et al.)

- a. Paul Graham essays
  - b. Wikipedia (English)
  - c. Shakespeare
  - d. [Maths textbook](#)
  - e. Something comparable in spirit but of your choice (do confirm with TA Ayush)
2. Learn the following models on XOR dataset (refer to Tensorflow Playground and generate the dataset on your own containing 200 training instances and 200 test instances) such that all these models achieve similar results (good). The definition of good is left subjective – but you would expect the classifier to capture the shape of the XOR function.
    - a. a MLP
    - b. MLP w/ L1 regularization (you may vary the penalty coefficient by choose the best one using a validation dataset)
    - c. MLP w/ L2 regularization (you may vary the penalty coefficient by choose the best one using a validation dataset)
    - d. learn logistic regression models on the same data with additional features (such as  $x_1 \times x_2$ ,  $x_1^2$ , etc.)Show the decision surface and comment on the plots obtained for different models. [2 marks]
  3. Using the [Mauna Lua CO2 dataset](#) (monthly) perform forecasting using an MLP and compare the results with that of MA (Moving Average) and ARMA (Auto Regressive Moving Average) models. Main setting: use previous “K” readings to predict next “T” reading. Example, if “K=3” and “T=1” then we use data from Jan, Feb, March and then predict the reading for April. Comment on why you observe such results. For MA or ARMA you can use any library or implement it from scratch. The choice of MLP is up to you. [2 marks]
  4. Train on MNIST dataset using an MLP. The original training dataset contains 60,000 images and test contains 10,000 images. If you are short on compute, use a stratified subset of a smaller number of images. But, the test set remains the same 10,000 images. Compare against RF and Logistic Regression models. The metrics can be: F1-score, confusion matrix. What do you observe? What all digits are commonly confused?

Let us assume your MLP has 30 neurons in first layer, 20 in second layer and then 10 finally for the output layer (corresponding to 10 classes). On the trained MLP, plot the t-SNE for the output from the layer containing 20 neurons for the 10 digits. Contrast this with the t-SNE for the same layer but for an untrained model. What do you conclude?

Now, use the trained MLP to predict on the Fashion-MNIST dataset. What do you observe? How do the embeddings (t-SNE viz for the second layer compare for MNIST and Fashion-MNIST images) [3 marks]