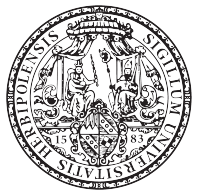

Machine Learning for Complex Networks
SoSe 2023

Prof. Dr. Ingo Scholtes
Chair of Informatics XV
University of Würzburg

Exercise Sheet 10

Published: July 5, 2023
Due: July 13, 2023
Total points: 10

Please upload your solutions to WueCampus as a scanned document (image format or pdf), a typesetted PDF document, and/or as a `jupyter` notebook.

1. **Convolutional Neural Network**

   (a) Implement a padding function that takes in a matrix of any size (representing a picture), the padding size and returns the padded matrix. 2P

   (b) Implement 2 different models and compare by visualising the loss function (all in the same plot). The data set to be used is the digit recognition MNIST dataset. Models to be compared: 2P

   a) Flatten all the pictures and feed them to an MLP model to classify the dataset. The MLP model should have 3 hidden layers of different sizes.

   b) Make use of Conv2d layer to implement a model with 2 different hidden layers, a pooling layer after each conv layer and lastly flatten the pictures to be feed to a linear layer.

   The choice of activation and optimization functions are left to the student. The number of epochs for each model is 300.

2. **Graph Convolutional Network**

   (a) Given the following neighborhood aggregation formula of GCN: 2P

   $$h_v^{l+1} = \sigma\big(W_l \sum_{u \in N(v)} \frac{h_u^{(l)}}{|N(v)|} + B_l h_v^l\big), \forall l \in \{0, \ldots, L-1\}$$

   derive the matrix formulation of $\sum_{u \in N(v)} \frac{h_u^{(l)}}{|N(v)|}$.

   (b) Generate a network of 150 nodes with 2 visible class distinctions. Create a GCN model with 2 conv layers, use softmax function to output the probabilities of classes. As a feature matrix for nodes generate: 2P

   a) A feature matrix of zeros,

   b) The identity matrix as a feature matrix,

   c) Use node2vec with $p = 1$ and $q = 4$ to get the node embeddings, and use the node embeddins as feature matrix.

   Train the model for 1000 epochs and plot the loss functions for each model in one plot. Comment on the differences of these models by checking the convergence of the loss function.

   (c) Use the latent space representation of the nodes from the second layer from one of the models mentioned above to visualise the feature maps. 2P