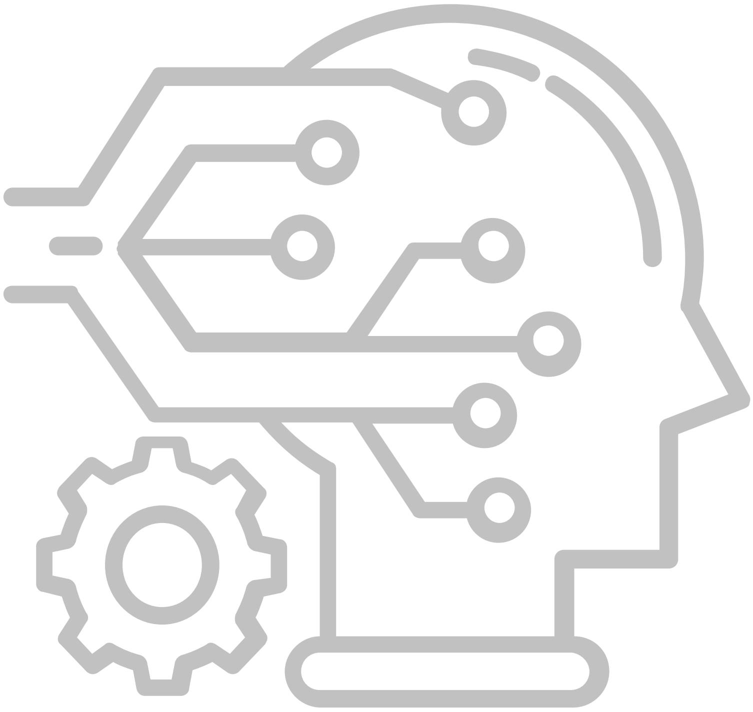
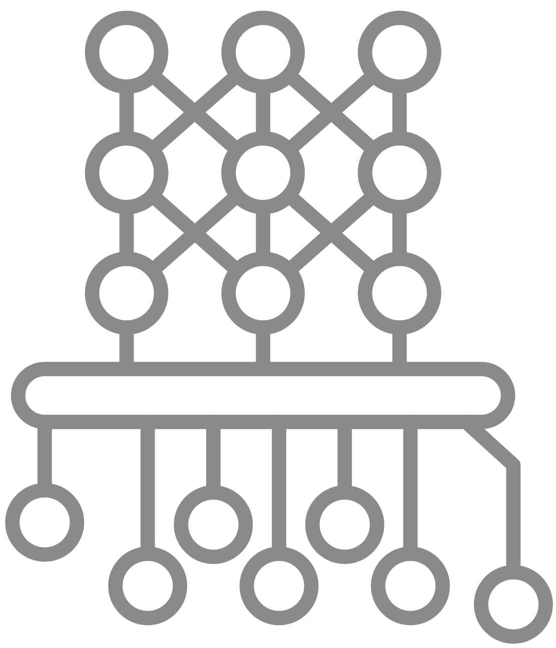


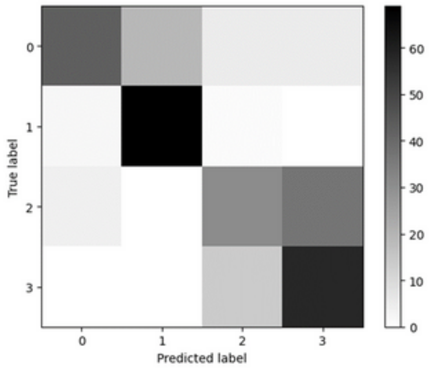
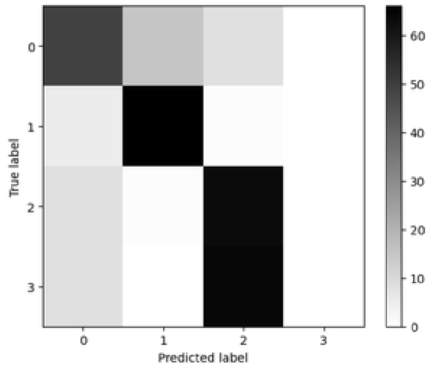
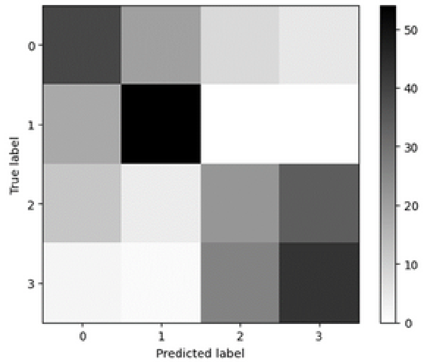
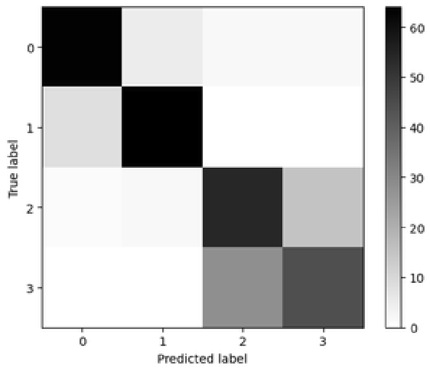
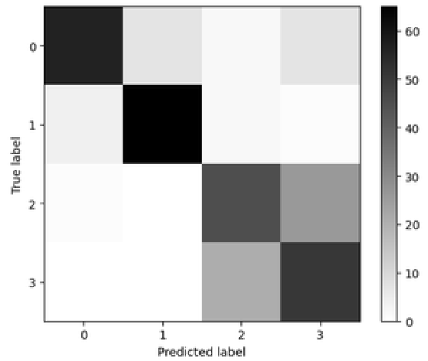
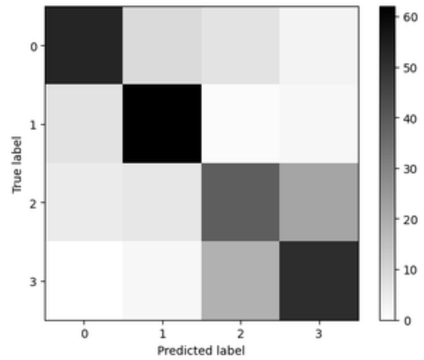
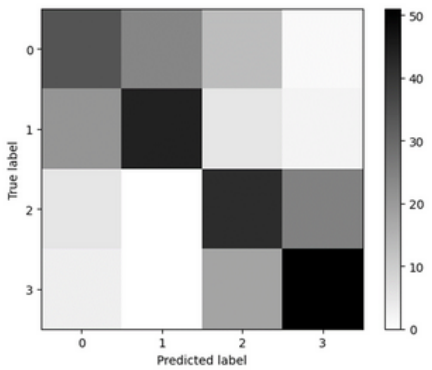
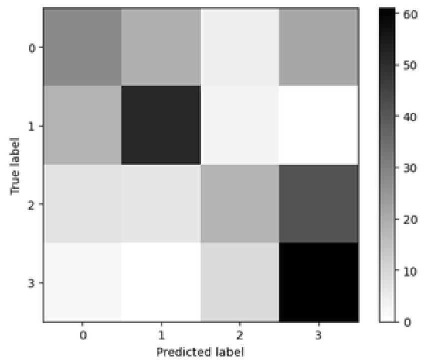
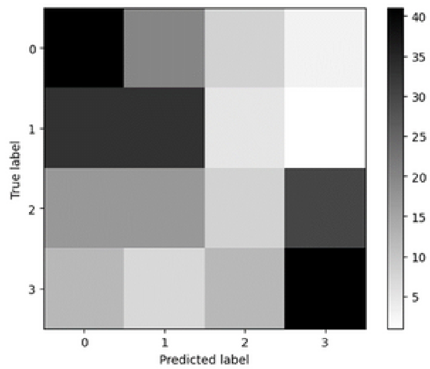
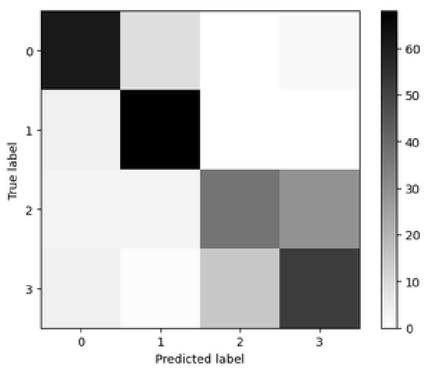
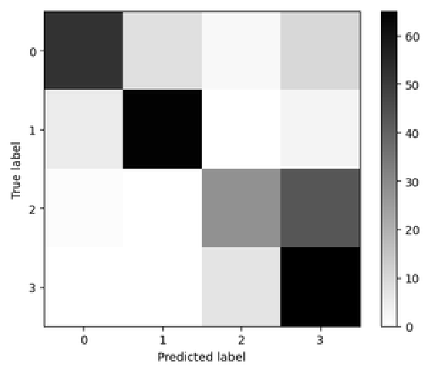
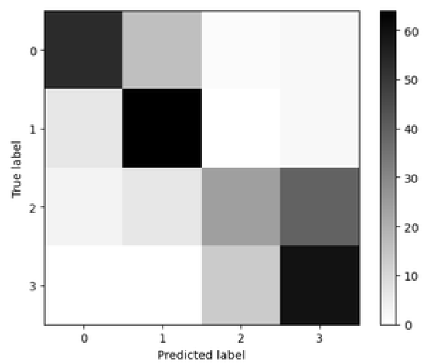
BCI COURSE

HW3 : Deep Learning for BCI



Part 1. Requirements

1.4.1 Show testing accuracies and plot confusion matrices

	EEGNet	SCCNet (or SCCNet_v2)	ShallowConvNet
Ind	 **Testing accuracy: 0.6979**	 **Testing accuracy: 0.6181**	 **Testing accuracy: 0.5486**
SI	 **Testing accuracy: 0.7812**	 **Testing accuracy: 0.7535**	 **Testing accuracy: 0.7118**
SD	 **Testing accuracy: 0.5938**	 **Testing accuracy: 0.5486**	 **Testing accuracy: 0.4271**
SI+FT	 **Testing accuracy: 0.7569**	 **Testing accuracy: 0.7257**	 **Testing accuracy: 0.6944**

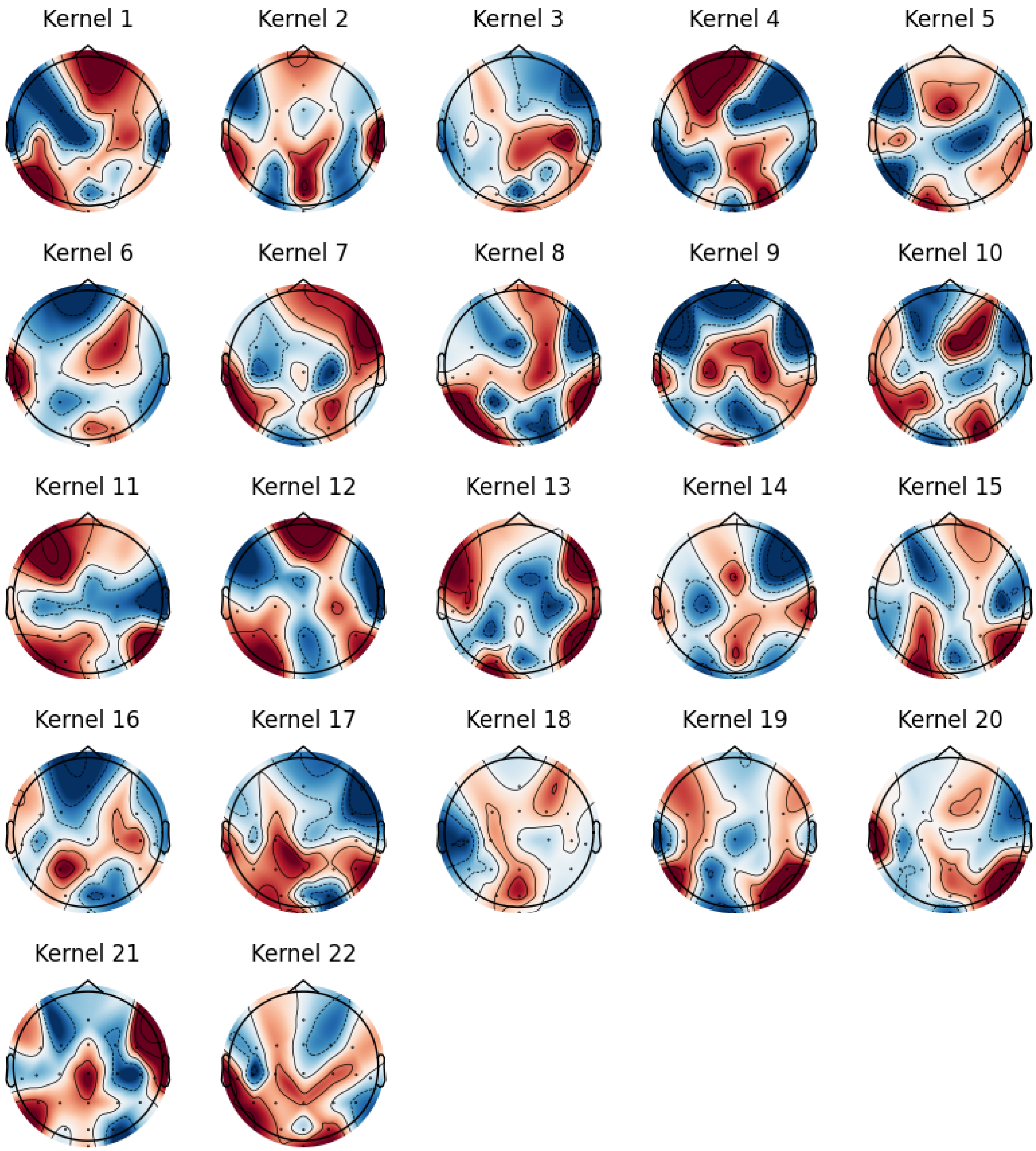
Part 1. Requirements

1.4.2 Fill in your hyper-parameter settings

	EEGNet	SCCNet (or SCCNet_v2)	ShallowConvNet
Ind	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"
SI	epochs = 250 batch_size =32 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"
SD	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"
SI+FT	epochs = 250 batch_size =32 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"	epochs = 200 batch_size =16 lr = 1e-3 subject_id = "01"

Part 1. Requirements

1.4.3 Obtain spatial kernel weights from the first convolutional layer of your SCCNet model trained with Ind scheme on subject 1, visualize the weights as topographic maps using the MNE package.



Part 2. Discussion

1.5.1 Pros and cons of the 3 CNN models and 4 training schemes

EEGNet

Pros:

1. Designed specifically for EEG data, so it incorporates some domain-specific knowledge.
2. Uses a separable convolutional layer, which can reduce the number of parameters and increase the efficiency of the model.
3. Has a relatively small number of parameters compared to other CNN models, which can make it easier to train.

Cons:

1. Limited model capacity due to its small size.
2. May not perform as well as other models on more complex EEG datasets.

ShallowConvNet

Pros:

1. Uses a simple architecture that is easy to understand and implement.
2. Has a relatively small number of parameters, which can make it easier to train.
3. Performs well on some EEG datasets.

Cons:

1. May not perform as well as other models on more complex EEG datasets.

SCCNet

Pros:

1. Uses a complex architecture that can capture more complex EEG patterns.
2. Has a larger number of parameters than other models, which can increase model capacity.
3. Performs well on some EEG datasets.

Cons:

1. May be more difficult to train due to its larger size.
2. May be more prone to overfitting due to its larger number of parameters.

Part 2. Discussion

1.5.1 Pros and cons of the 3 CNN models and 4 training schemes

Individual (Ind) Training Scheme

Pros:

1. Can result in a model that is optimized for a specific individual's EEG data.
2. May perform better than other training schemes on EEG datasets with high inter-subject variability.

Cons:

1. Requires a separate model to be trained for each individual, which can be time-consuming and computationally expensive.

Subject Independent (SI) Training Scheme

Pros:

1. Can result in a more generalizable model that can be applied to new subjects.
2. Can be less computationally expensive than the Ind training scheme since only one model needs to be trained.

Cons:

1. May not perform as well as other training schemes on EEG datasets with high inter-subject variability.

Subject Dependent (SD) Training Scheme

Pros:

1. Can result in a model that is optimized for a specific subject's EEG data.
2. May perform better than the SI training scheme on EEG datasets with low inter-subject variability.

Cons:

1. Requires a separate model to be trained for each subject, which can be time-consuming and computationally expensive.

Subject Independent + Fine Tuning (SI+FT) Training Scheme

Pros:

1. Combines the advantages of the SI training scheme with the ability to fine-tune the model for individual subjects.
2. Can result in a more generalizable model that can be fine-tuned for individual subjects as needed.

Cons:

1. Requires additional training time and computational resources compared to the SI training scheme.

Part 2. Discussion

1.5.2 Your observations regarding the models

Structures

EEGNet, ShallowConvNet, and SCCNet are all designed specifically for EEG data, but they have different architectures. EEGNet has a deeper architecture with separable convolutional layers, while ShallowConvNet has a simpler architecture with fewer layers. SCCNet has a more complex architecture that combines convolutional and recurrent layers. Overall, EEGNet and SCCNet have more complex structures than ShallowConvNet.

Batch size

The optimal batch size for each model may differ depending on the specific dataset and hardware used. However, in general, EEGNet and SCCNet can handle larger batch sizes than ShallowConvNet due to their deeper architectures and parallelization techniques.

Learning rate

The optimal learning rate for each model may also differ depending on the dataset and other hyperparameters. In general, however, EEGNet and ShallowConvNet may require smaller learning rates than SCCNet due to their simpler architectures. SCCNet's more complex architecture may require a higher learning rate to effectively update all of the parameters.

Other hyperparameters

Other hyperparameters, such as the number of filters, filter sizes, dropout rates, and activation functions, can also impact the performance of the models. Tuning these hyperparameters may require experimentation and can vary depending on the dataset and task at hand.

Overall

EEGNet and SCCNet may be better suited for more complex EEG classification tasks that require more layers and higher-level representations, while ShallowConvNet may be better suited for simpler tasks or situations where computational resources are limited. However, the optimal choice of model will depend on the specific dataset, task, and available resources.

Part 2. Discussion

1.5.3 Difficulties you encountered in this homework.

Throughout this homework, I faced several challenges related to EEG data and model development. Firstly, my limited knowledge of EEG made it difficult for me to interpret the data and understand its characteristics, which in turn made it challenging to pre-process the data and perform the necessary analysis. This lack of familiarity with the data also caused issues when calculating the correct size, leading to errors during the training process. I found it frustrating that I was not able to gain a more intuitive understanding of the data, which made troubleshooting issues more difficult.

Additionally, while working with the SCCNet model, I encountered several issues related to its implementation. Although the model followed the paper frame, I had difficulty framing it correctly, which made it challenging to understand the model architecture and implement it correctly. This challenge required me to put in additional effort and spend more time studying the model before I was able to implement it successfully.

Overall, these challenges made the homework more difficult than I had anticipated, and it reminded me of the importance of having a deep understanding of the data and the model architecture. Although I encountered several roadblocks along the way, I am grateful for the opportunity to learn from these challenges and develop my skills in EEG data analysis and deep learning.

Part 2. Discussion

1.5.4 For models trained with different subjects, what are the possible reasons for the difference in model performance

There are several possible reasons why models trained with different subjects may have different performances.

- **Variability in EEG signals**： EEG signals can vary significantly between different individuals due to factors such as differences in brain structure, head size and shape, and electrode placement. As a result, models trained on EEG data from one subject may not generalize well to data from another subject.

- **Variability in task performance**： Even when the same task is performed by different individuals, there may be differences in the way they approach the task or in their level of engagement, which can affect the EEG signals recorded. This can result in variability in the data that the model is trained on, leading to differences in performance on test data from different subjects.

- **Differences in data quality**： EEG data can be affected by a variety of noise sources, such as muscle activity, eye movements, and environmental interference. Differences in the quality of the data collected from different subjects can affect the performance of the model trained on that data.

- **Overfitting**： Models trained on data from one subject may overfit to that specific subject's data, resulting in poor performance on test data from other subjects. This can occur if the model is too complex and is able to memorize the specific patterns in the training data, rather than learning generalizable features.

- **Model architecture**： Different model architectures may perform differently on data from different subjects, depending on the complexity of the task and the nature of the EEG signals. For example, a model with a more complex architecture may be able to capture subtle differences in EEG signals that are specific to one subject, but may not generalize well to other subjects.