Module 6

Hyperparameter Tuning and Feature Engineering

▼ What is rescaling?

- Rescaling data simply means translating the values in some way to change extreme values in the data and intermediate values are moved in a way that is consistent.
- Rescaling can be fixed, meaning the rule you are scaling by uses a fixed value, or it can be standardized, compressing values based on the spread as measured by the variance.
- In the preprocessing module of scikit-learn the scale( ) function can be used for rescaling feature values in this way.

▼ What is extrapolation?

Extrapolation is a sort of estimation of a variable's value beyond the initial observation range based on its relationship with another variable.

▼ What is a feature?

a particular attribute, property, or data type that is both measurable as well as characteristic of the objective function that you are trying to approximate with our model

▼ What is feature selection?

Feature selection drops columns from a feature array, rather than rows. This way, you can remove redundant and irrelevant features.

▼ What are wrapper methods (feature selection)?

- Methods such as Recursive Feature Eliminator (RFE) in the feature_selection module of scikit-learn are known as wrapper feature selection methods.
- These methods use different subsets of the data to train multiple models, from which the model that provides the best performance is selected.

▼ What are filter methods (feature selection)?

Filter methods rely on domain knowledge and knowledge of the datatypes used to apply statistical techniques to reduce redundant predictive features and prevent problems with collinearity.

▼ Why is Lasso regression used?

- The lasso regression allows you to shrink or regularize these coefficients to avoid overfitting and make them work better on different datasets.
- This type of regression is used when the dataset shows high multicollinearity or when you want to automate variable elimination and feature selection.

▼ How can decision trees be filtered of features?

by limiting the search depth

▼ What is feature construction?

- Feature construction is a process that creates intermediate features in a dataset from the original descriptors.
- The goal is to create more efficient features for a machine learning to draw from in order to approximate a function accurately.

▼ What is feature interaction?

Features also often interact, meaning the inclusion of one feature with another causes the model to output a particular result that may be different when only one of the features is present.

▼ What are some techniques for automatic feature selection?

- Genetic programming

- ANNs designed with domain knowledge can be built and trained as feature extractors to produce more diversified and highly informative features

▼ What is feature extraction?

any dimensionality reduction process that extracts and reduces features into more manageable feature groups for easier processing

▼ What is feature coding?

a technique used in image classification used to obtain representative codings of feature descriptors using an unsupervised trained dictionary

▼ What are hyperparameters?

- Hyperparameters, sometimes called tuning parameters, are important parameters which the data cannot estimate.
- They represent high-level properties of a model.
- Note: you will need to have these parameters identified before the initialization of model training.

▼ What are the two types of hyperparameters?

- Model hyperparameters: influence the performance and accuracy of the model and which are considered when undergoing model selection
- Algorithm hyperparameters: affect the speed and quality of the training process.

▼ What are some examples of hyperparameters?

- the leaves or depth of a search tree
- the learning rate of a model
- the number of clusters in a k-means clustering model
- alpha in a lasso regression model
- K in K-NN models
- the C penalty of SVM classifiers

▼ What is the purpose of cross-validation?

- To test the ability of a machine learning model to predict new data.
- When tuning hyperparameters, an automatic process is set up to evaluate the hyperparameters one by one and compare the accuracy of the resulting model's outputs, inference speed, and hyperparameter interaction.

▼ What is K-fold cross validation?

- When validating your test set, you divide your sample data into k subsets and train your model on k-1 subsets, reserving the final subset for a test.
- You repeat this process k times, adjusting your hyperparameters each iteration until you find a model with suitable accuracy.

▼ What is GridSearch?

GridSearch-based hyperparameter tuning uses a dictionary or grid of *n* dimensions of possible hyperparameters and values as input and search the entire grid for a configuration that results in a model that provides an ideal test output.

▼ What is RandomSearch?

Random search is similar to GridSearch except that the search space begins at a random spot on the grid.

▼ What is Orthogonal Array Tuning?

- A hyperparameter tuning technique uses an array with entries arranged in a specific way.
- Specifically, for every selection of two different columns of the table, all ordered 2-tuples of the elements appear for the same number of times, called a factor-level table, that runs models based on hyperparameters provided in a corresponding Orthogonal Array Tuning table.