Slim Ben Salah

Machine Learning: Modeling

All You Need to Know!

Modeling & Machine Learning: 1/4

Model Types

- **Classification models:** These models learn from labeled data to categorize new instances into predefined classes or categories.

- **Regression models:** By analyzing the relationship between variables, regression models make continuous predictions. They uncover underlying patterns between input variables and a target variable, enabling estimation for new data.

- **Clustering models:** Clustering models group similar data points together based on their characteristics and patterns. They help discover hidden structures or clusters within the data.

- **Neural Networks:** Inspired by the human brain, neural networks consist of interconnected layers of nodes (neurons). They process and transform input data, enabling complex learning and pattern recognition tasks.

Modeling & Machine Learning: 2/4

Popular ML Models

- **Classification:** Logistic Regression.
  Decision Trees.
  Support Vector Machines (SVM).
- **Regression:** Linear Regression.
  Random Forest Regression.
  Gradient Boosting Regression.
- **Clustering:** K-Means.
  Hierarchical Clustering.
  DBSCAN.
- **Neural Networks:** Feedforward Neural Networks.
  Convolutional Neural Networks (CNNs).
  Recurrent Neural Networks (RNNs).

Slim Ben Salah

Modeling & Machine Learning: 3/4

Hyperparameter Tuning

Hyperparameter tuning involves finding the optimal values for the hyperparameters of a machine learning model. Hyperparameters are configuration settings whose values are used to control the learning process. These parameters control the behavior and performance of the model. Tuning these hyperparameters is important as it can significantly impact the model's performance and ability to generalize well to new data.

- **Grid Search:** Exhaustively search through a predefined set of hyperparameter values to find the best combination.

- **Random Search:** Randomly sample hyperparameter values from a defined range to explore different combinations.

- **Bayesian Optimization:** Utilize a probabilistic model to intelligently select hyperparameter values based on previous evaluations.

Swipe to Learn

Slim Ben Salah

Modeling & Machine Learning: 4/4

Overfitting & Underfitting

• **Overfitting** occurs when a model becomes too complex and learns noise or irrelevant patterns from the training data, resulting in poor generalization on unseen data. Signs of overfitting include high accuracy on the training data but low performance on the validation or test data, as well as a significant difference between the training and validation performance.
To address overfitting, techniques like regularization can be employed since it adds a penalty term to the model's loss function, discouraging complex patterns and promoting generalization.

• **Underfitting** happens when a model is too simplistic and fails to capture the underlying patterns in the data, leading to low accuracy and poor performance. Signs of underfitting include low accuracy on both the training and validation data, as well as a small gap between the training and validation performance.
To address underfitting, increasing the model's complexity by adding more layers to a neural network or using more advanced models, such as deep neural networks or ensemble methods, can help capture the complexity of the data and improve performance.