HW-Topic-11
Data Acquisition, Modeling and Analysis: Big Data Analytics

Submitted By – **Sudhanshu Kakkar**
CWID – **20036779**

Feature Selection

CORE IDEA

- Choosing the most important input features for a model.

- Helps the model focus on what matters.

- Removes noise and makes learning easier.

GOAL

- Reduce dimensionality.

- Improve model accuracy.

TYPES

Type	Speed	Accuracy	Model Used	Examples
Filter	Fast	Low-Medium	No	Correlation, Chi Square
Wrapper	Slow	High	Yes (many times)	Forward Selection
Embedded	Medium	Medium-High	Yes (during training)	Lasso

Advantages

- Faster training and prediction.

- Less overfitting.

- Cleaner and simpler models.

- Better accuracy when data has noise.

Disadvantages

- May remove useful information

- Hard to decide the right number of features

- Wrapper methods are slow

- Can be computationally heavy