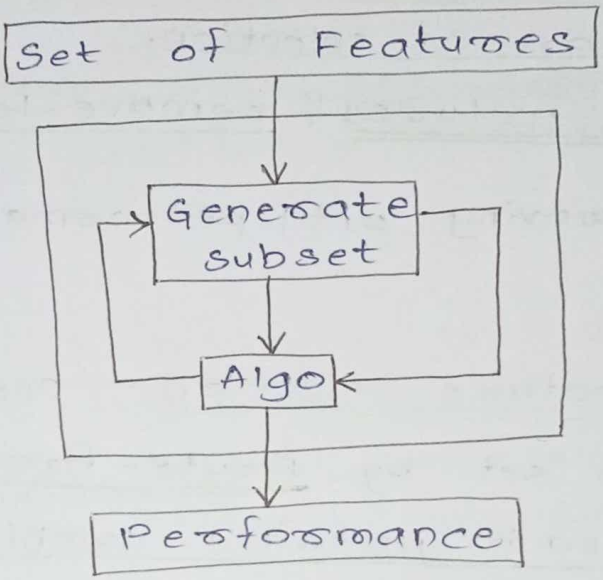
2) Wrapper Methods :-

↳ Find subset of features that optimizes the performance of ML model.

↳ create different combinations of subsets from feature sets, train the model, evaluate the model performance & compare. Then, select model with best performance.



- Steps in Wrapper Methods:

① Start with initial set of features (empty set / all features).

② Train model using selected feature subset.

③ Evaluate model performance using chosen metrics - accuracy, F1-score / cross-validation.

④ Modify subset based on result (ex. add, remove or swap features).

⑤ Repeate steps 2 to 4 until stopping criterion is meet (ex. meet num. of features / meet highest improvement).

- Some Techniques :-

① Forward Selection Method - Begin with empty set.

↳ Iteratively add features & evaluate whether performance improve / not.

↳ Add until go on improve performance.

② Backward Elimination — Iterative approach.

↳ Opposite of foward selection.

↳ Start with all features & remove least significant features.

↳ Continue improving until performance improve.

③ Exhaustive Feature selection — One of the best.

↳ Evaluate each set by Brute-force.

↳ Try & make each possible combination of features & return best performance subset.

④ Recursive Feature Elimination —

↳ Recursive Greedy Optimization method.

↳ Select features by recursively taking smaller & smaller feature subsets.

↳ Now, estimator is trained with each subset & each feature importance is determined by coef_attribute / through feature_importance_attribute.

- **Advantages of Wrapper Methods —**

① **Model-centric** — consider **specific ML model** & **select most relevant features** according to model performance.

② **Feature Interaction consideration** — can capture **interactions between features**, unlike filter methods.

③ **Optimization** — Find **best subset of features** by optimizing specific performance metric.

- **Limitations of Wrapper Methods —**

① **Computational cost** — computationally expensive, especially **large datasets** / **exhaustive search**.

② **Overfitting Risk** — overfitting risk, especially for **small datasets**.

③ **Model selection** — ML Model influence results. **Different models** ⇒ **different optimal features**.