```
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐
│    Step 1:      │ ───> │    Step 2:      │ ───> │    Step 3:      │
│ Data Collection │      │ EDA (Exploratory│      │ Feature Analysis│
│                 │      │ data Analysis)  │      │ Feature Selection│
│                 │      │                 │      │ Feature Extraction│
└─────────────────┘      └─────────────────┘      └─────────────────┘
                                                          │
                                                          v
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐
│    Step 4:      │ <─── │    Step 5:      │ <─── │    Step 6:      │
│ Preprocessing   │      │  Model Train    │      │ Evaluate (Cross │
│                 │      │                 │      │ Validation,     │
│                 │      │                 │      │ Dividing train  │
│                 │      │                 │      │ and test datasets)│
└─────────────────┘      └─────────────────┘      └─────────────────┘
   │
   v
┌─────────────────┐      ┌─────────────────┐
│    Step 7:      │ ───> │    Step 8:      │
│  Improvement    │      │ Using Algorithm │
│ Different ways   │      │ using algorithm │
│ like some-other  │      │ to predict      │
│ preprocessing    │      │ the output value│
│ techniques,      │      │ etc.            │
│ using another    │      │                 │
│ model etc.       │      │                 │
└─────────────────┘      └─────────────────┘
```

Exploratory data analysis (EDA) is used by data scientists to analyze and investigate data sets and summarize their main characteristics

Feature Selection is **the method of reducing the input variable to your model by using only relevant data and getting rid of noise in data**

Feature Selection Techniques

Supervised Feature Selection

Filters method
→ Missing value
→ Information gain
→ Chi-square Test
→ Fisher's Score

Embedded method
→ Regularization L1, L2
→ Random forest Importance

Wrappers method
→ Forward Feature Selction
→ Backward Feature Selection
→ Exhaustive Feature Selection
→ Recursive Feature Elimination

Unsupervised Feature Selection