## Sequential forward feature selection

- X= {f1, f2, f3, f4, f5}, y, select 3 features
- Initialisation:
  - Sel = {}
  - Rem = {f1, f2, f3, f4, f5}
  - Start loop
- Iteration 1: currently Sel = {}, Rem = {f1, f2, f3, f4, f5}
  - score(f1) = 0.5, score(f2)=0.6, score(f3) = 0.2, score(f4) = 0.3, score(f5) = 0.7
  - Select f5
  - Sel = {f5}
  - Rem = {f1, f2, f3, f4}

## Sequential forward feature selection

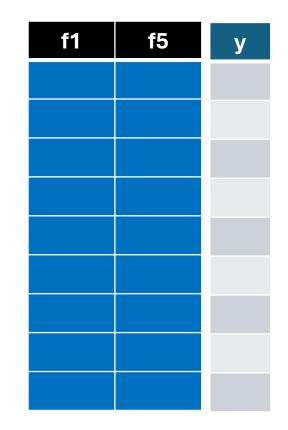
- Iteration 2: currently Sel = {f5}, Rem = {f1, f2, f3, f4}
  - score({f1, f5}) = 0.7, score({f2, f5})=0.2, score({f3, f5}) = 0.1, score({f4, f5}) = 0.2
  - Select f1
  - Sel = {f5, f1}
  - Rem = { f2, f3, f4}
- Iteration 3: currently Sel = {f5, f1}, Rem = {f2, f3, f4}
  - score({f2, f5, f1}) = 0.8, score({f3, f5, f1})=0.1, score({f4, f5, f1}) = 0.05
  - Select f2
  - Sel = {f5, f1, f2}
  - Rem = {f3, f4}
- Stop since 3 features are selected, output Sel = {f5, f1, f2}

### How to evaluate {f1, f5}?

### Training set

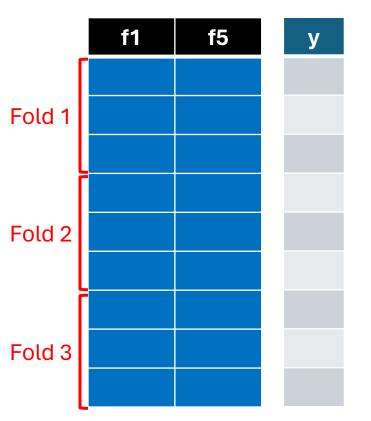
|        | f1 | f2 | f3 | f4 | f5 | У |
|--------|----|----|----|----|----|---|
| Fold 1 |    |    |    |    |    |   |
|        |    |    |    |    |    |   |
| Fold 2 |    |    |    |    |    |   |
|        |    |    |    |    |    |   |
| Fold 3 |    |    |    |    |    |   |
|        |    |    |    |    |    |   |
|        |    |    |    |    |    |   |

#### **Transformed Training set**



# How to evaluate {f1, f5}?

Transformed Training set



- Classifier clf
  - Fold 1 as sub\_testing, Fold 2+Fold 3 as sub\_training, train clf on Fold 2+Fold 3 and test clf on Fold 1 -> acc1
  - Fold 2 as sub\_testing, Fold 1+Fold 3 as sub\_training -> train clf on Fold 1+Fold 3 and test clf on Fold 2 -> acc2
  - Fold 3 as sub\_testing, Fold 1+Fold 2 as sub\_training -> train clf on Fold 1+Fold 2 and test clf on Fold 3 -> acc3
  - Ave\_acc = (acc1+acc2+acc3)/3