

Sequential forward feature selection

- **$X = \{f_1, f_2, f_3, f_4, f_5\}$, y , select 3 features**
- Initialisation:
 - $Sel = \{\}$
 - $Rem = \{f_1, f_2, f_3, f_4, f_5\}$
 - Start loop
- Iteration 1: currently $Sel = \{\}$, $Rem = \{f_1, f_2, f_3, f_4, f_5\}$
 - $score(f_1) = 0.5$, $score(f_2) = 0.6$, $score(f_3) = 0.2$, $score(f_4) = 0.3$, $score(f_5) = 0.7$
 - Select f_5
 - $Sel = \{f_5\}$
 - $Rem = \{f_1, f_2, f_3, f_4\}$

Sequential forward feature selection

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

How to evaluate $\{f1, f5\}$?

Transformed Training set

	f1	f5	y
Fold 1			
Fold 2			
Fold 3			

- 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**
- $\text{Ave_acc} = (\text{acc1} + \text{acc2} + \text{acc3}) / 3$