

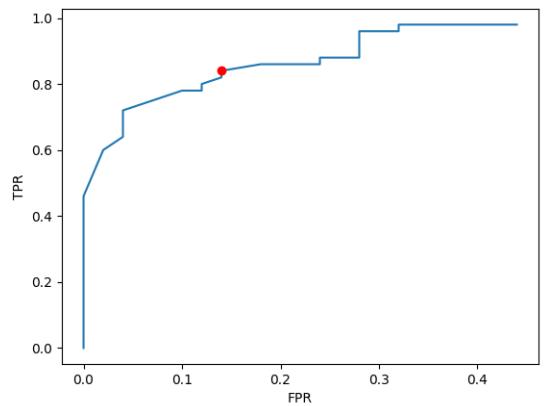
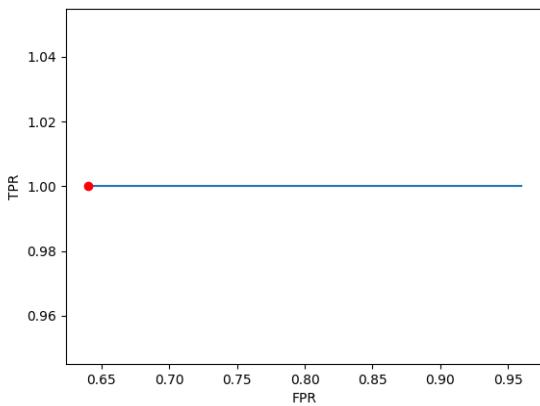
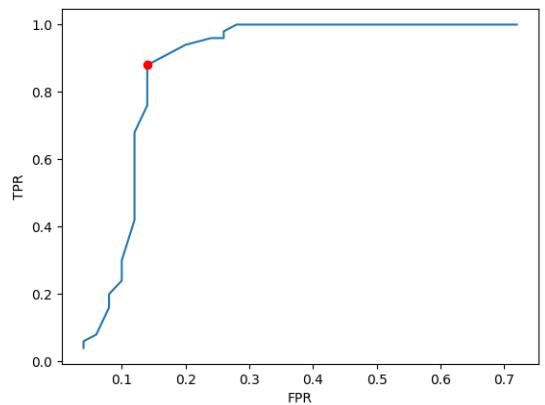
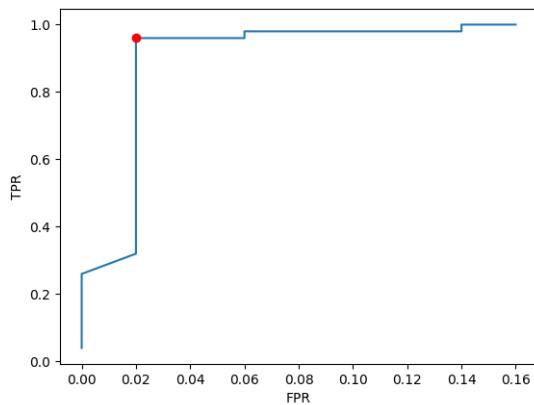
# Strojové učení - Report

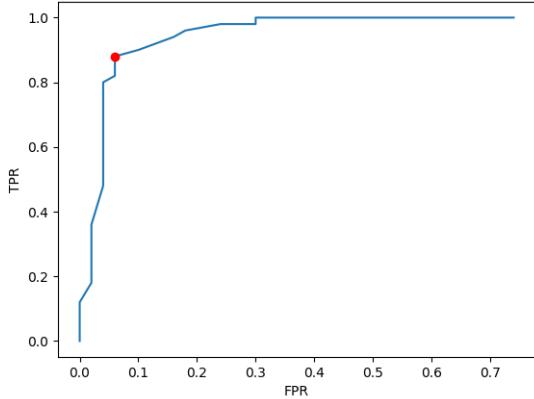
Jonáš Holub

June 4, 2021

## 1 Výběr vhodného parametru

K výběru vhodného parametru využijí ROC křivky pro daný klasifikátor. Parametr  $\alpha$  vyberu podle euklidovské vzdálenosti od ideálního klasifikátoru - bod  $[0, 1]$ .





Zvolím tedy klasifikátor C5 s parametrem  $\alpha_{16}$ .

## 2 Přísně tajné!

Pro požadavky zadání je parametr  $\alpha$  optimální, když je minimální počet falešně pozitivních případů. Pokud je takovýchto minim více, pak vyberu ten parametr, pro který nabývá počet správně klasifikovaných případů maxima. (Tedy vyberu ten parametr, pro který je vzdálenost od bodu  $[0, 1]$  na ROC křivce minimální)

Zvolím tedy klasifikátor C4 s parametrem  $\alpha_{11}$ .

## 3 Hlavně bezpečně

Klasifikátory budu porovnávat na základě vzdáleností od bodu  $[0, 1]$  na ROC křivce. Menší vzdálenost odpovídá lepšímu klasifikátoru.

Funkce by tedy mohla vypadat nějak takto:

```
def compare(c, alpha, c6):
    x1 = c.TPR[alpha]
    y1 = c.FPR[alpha]
    c6_TPR = TPR(c6) #list s true positive rates pro každý alfa
    c6_FPR = FPR(c6) #list s false positive rates pro každý alfa

    FPRmin = np.where(c6_FPR == c6_FPR.min())[0]
    i = np.where(c6_TPR == np.max(c6_TPR[FPRmin]))[0]

    x2 = c6_TPR[i]
    y2 = c6_FPR[i]

    dist = lambda x,y: np.sqrt(x**2 + (1-y)**2)
```

```
if dist(x1,y1) > dist(x2,y2):
    return True
else:
    return False
```