

11. Úkol na předmět Mnohorozměrná analýza

Jaromír Macoun

Květen 2022

1 Úvod

Zabýváme se v tomto úkolu daty *bankruptcy*, které obsahují záznamy o bankrotu a dále dvě numerické charakteristiky dané instituce. Přesněji řečeno data obsahují 84 pozorování o institucích, zdali zkrachovali, dále jejich profit a leverage.

2 Diskriminační analýza

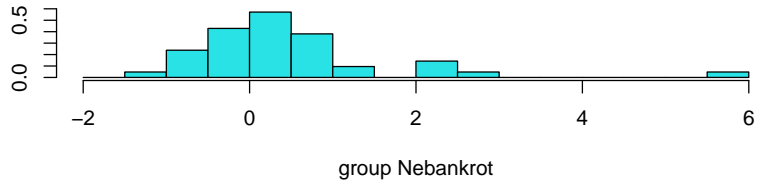
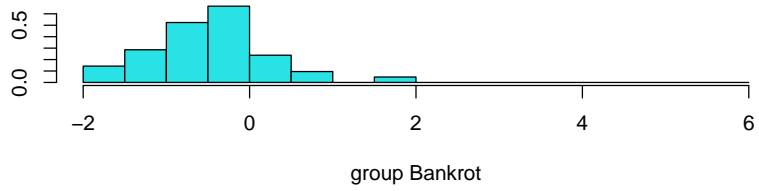
2.1 Lineární diskriminační analýza

Pomocí programu R byla provedena LDA metoda na data. Na základě obrázku č. 1 vidíme, že histogramy pro jednotlivé skupiny se soustřeďují okolo nuly. Pro Bankrot je histogram spíše vlevo od nuly, naopak pro nebankrot vpravo od nuly. Nicméně tyto rozdíly jsou velmi malé.

Na obrázku č. 2 je přímo znázorněna dělicí přímka a červeně o značené špatně predikované hodnoty. Dále je vidět, že zhruba 27% případech byla chybná predikce.

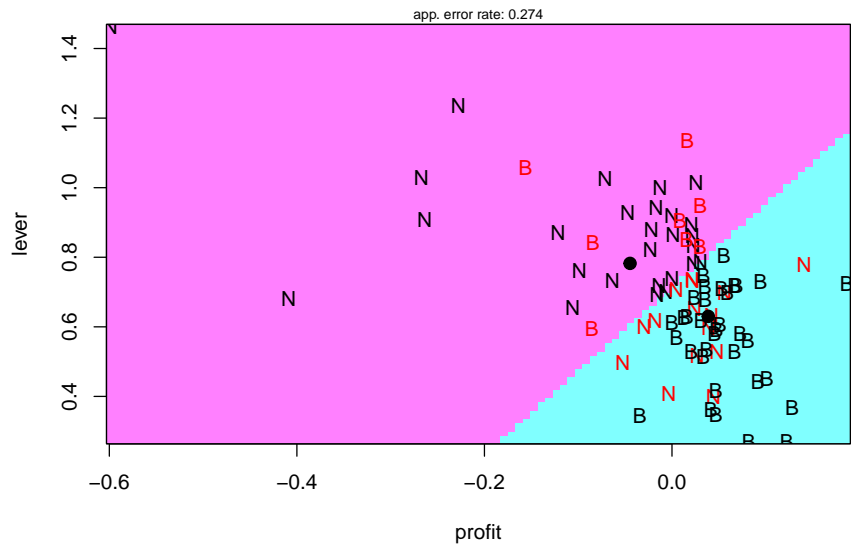
2.2 Kvadratická diskriminační analýza

Jako alternativní metodu k zjišťování zdali instituce zbankrotovala jsem si vybral kvadratickou diskriminační analýzu. Její výsledky jsou shrnuty na obrázku č. 3. Jak je vidět oproti LDA metodě jsme nedosáhli zlepšení co se týče app. error rate ten se zvýšil na téměř 40%. Jediné "zlepšení"by mohlo nastat v potenciální chybě prvního druhu. Pokud chápeme jako horší druh chyby, že predikujeme bankrot nicméně instituce nebankrotuje, tak v tomto ohledu došlo ke zlepšení. V LDA jsem v 8mi případech predikovali bankrot ale instituce nebankrotovala, zato v QDA metodě se tento počet snížil na 4.

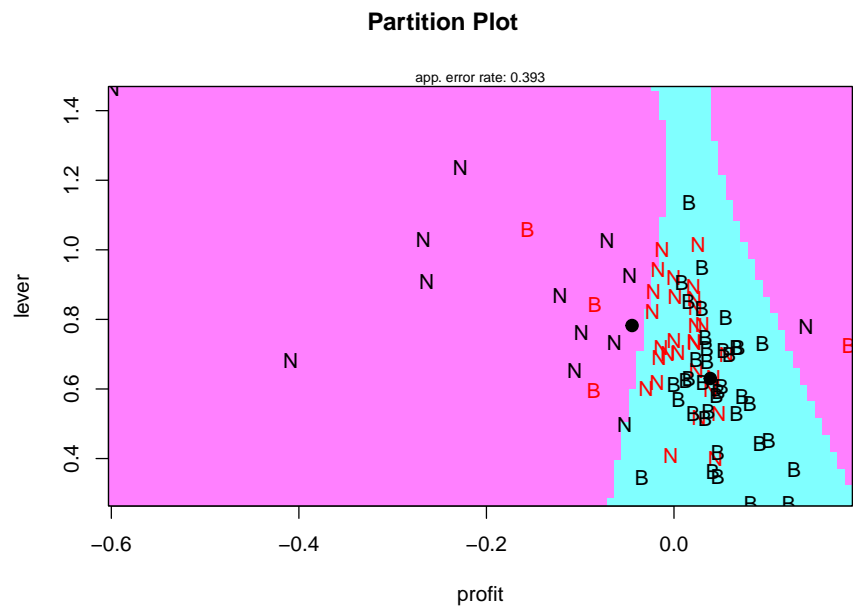


Obrázek 1

Partition Plot



Obrázek 2



Obrázek 3