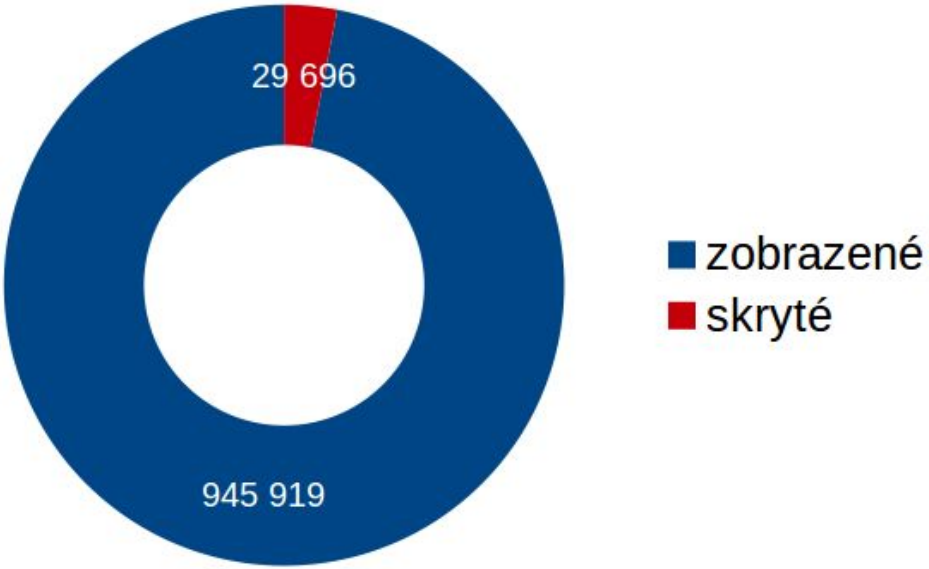
Klasifikácia príspevkov v diskusiách

Andrej Švec

moderateIT

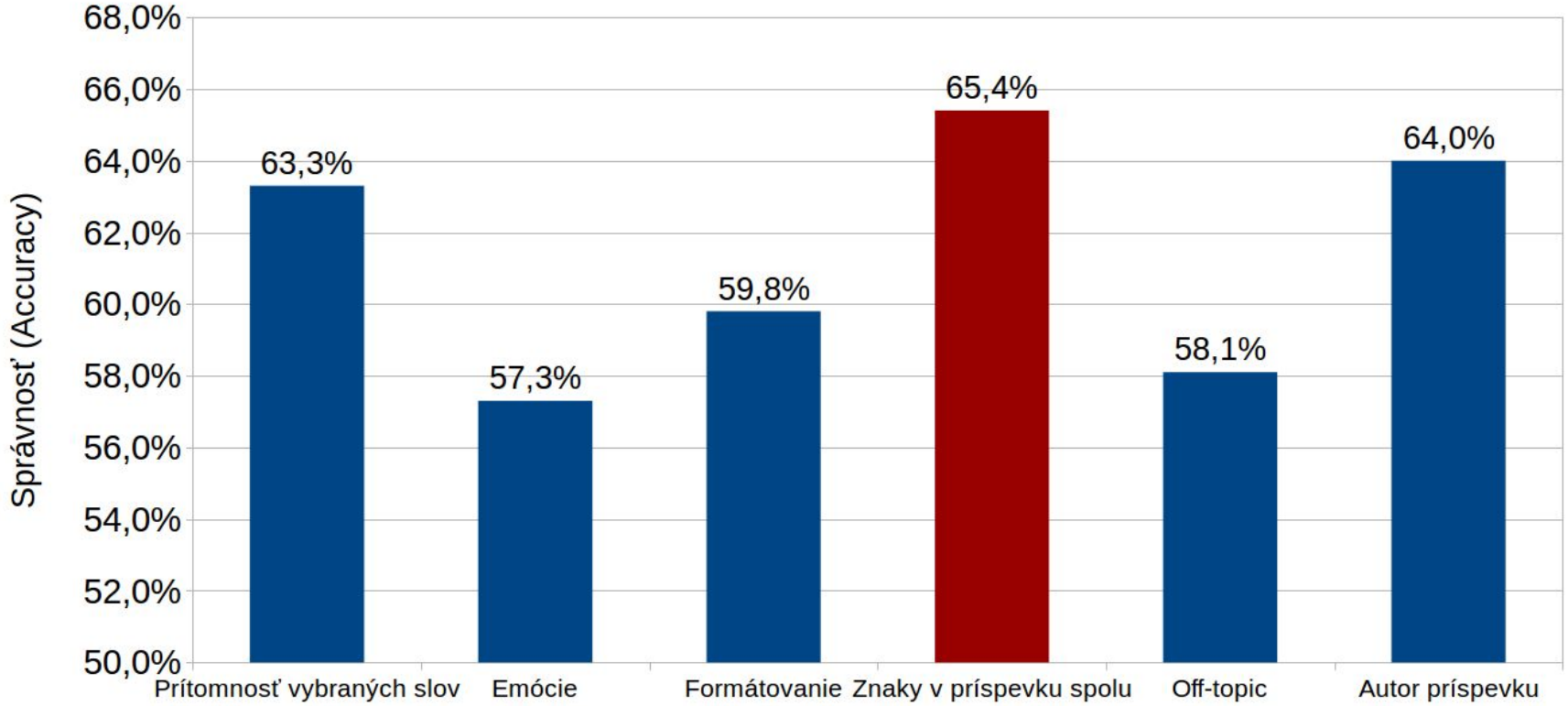
STU FIIT

Dataset zo SMEsk z roku 2015

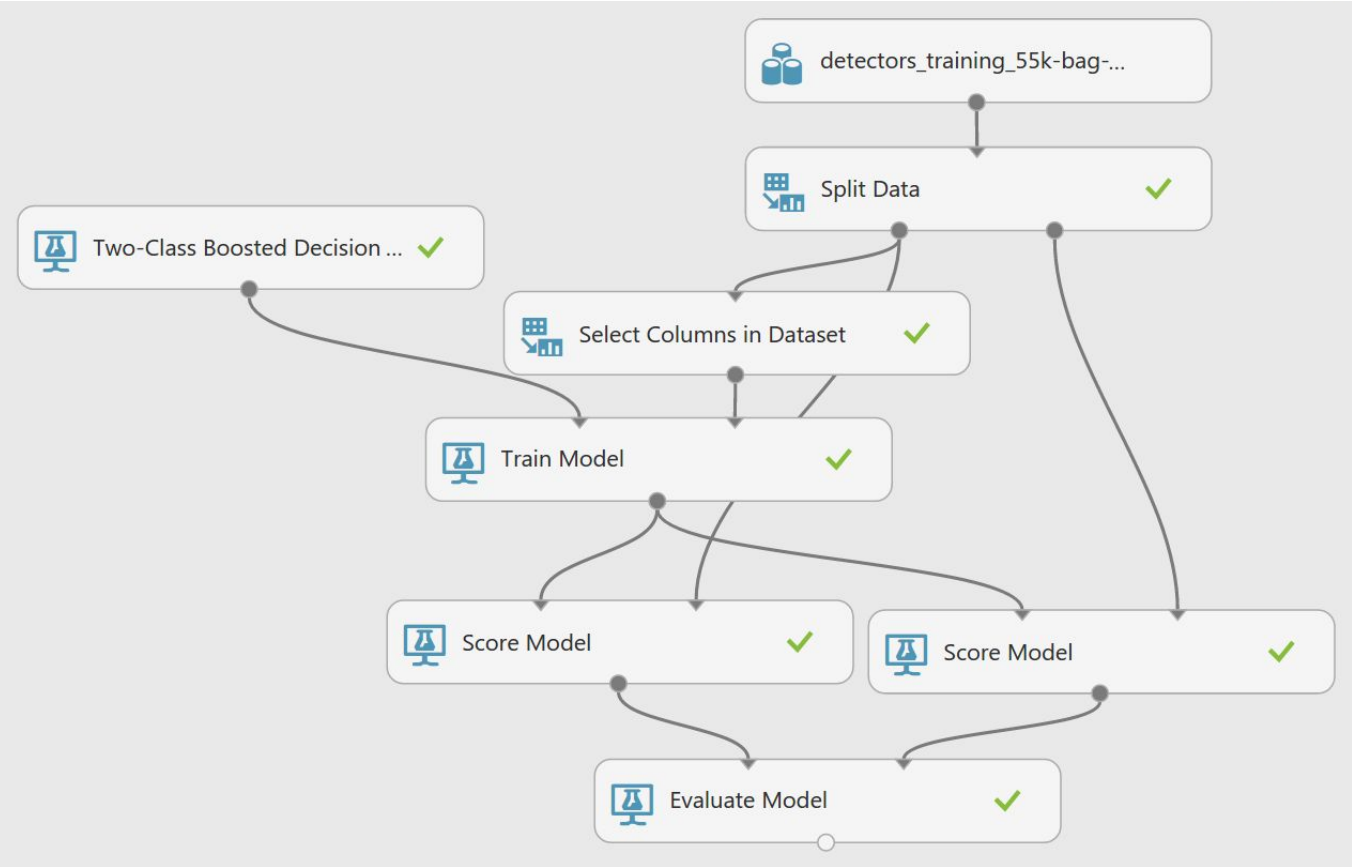


hodnotenie >13

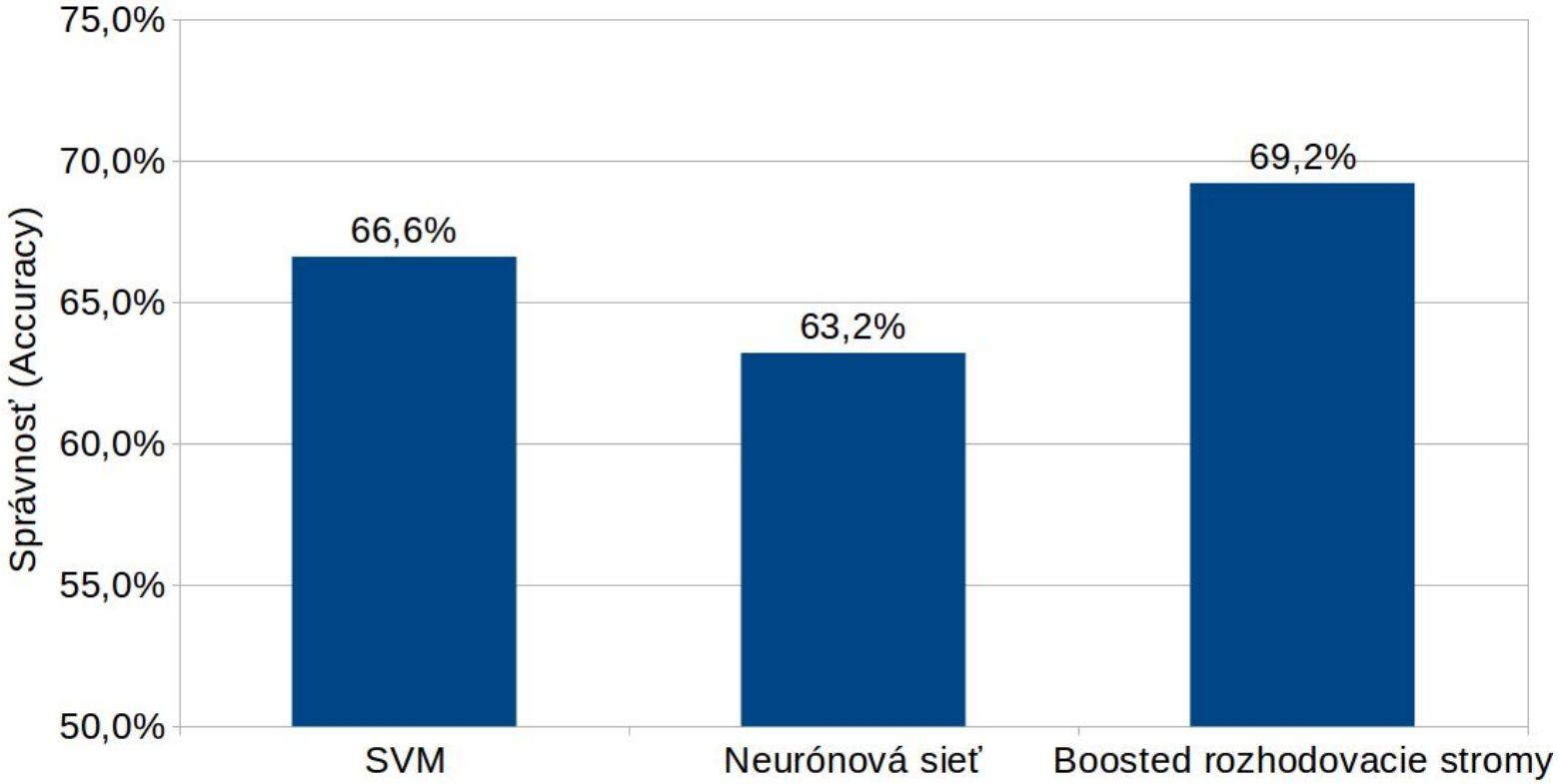
Klasifikácia na základe extrahovaných znakov



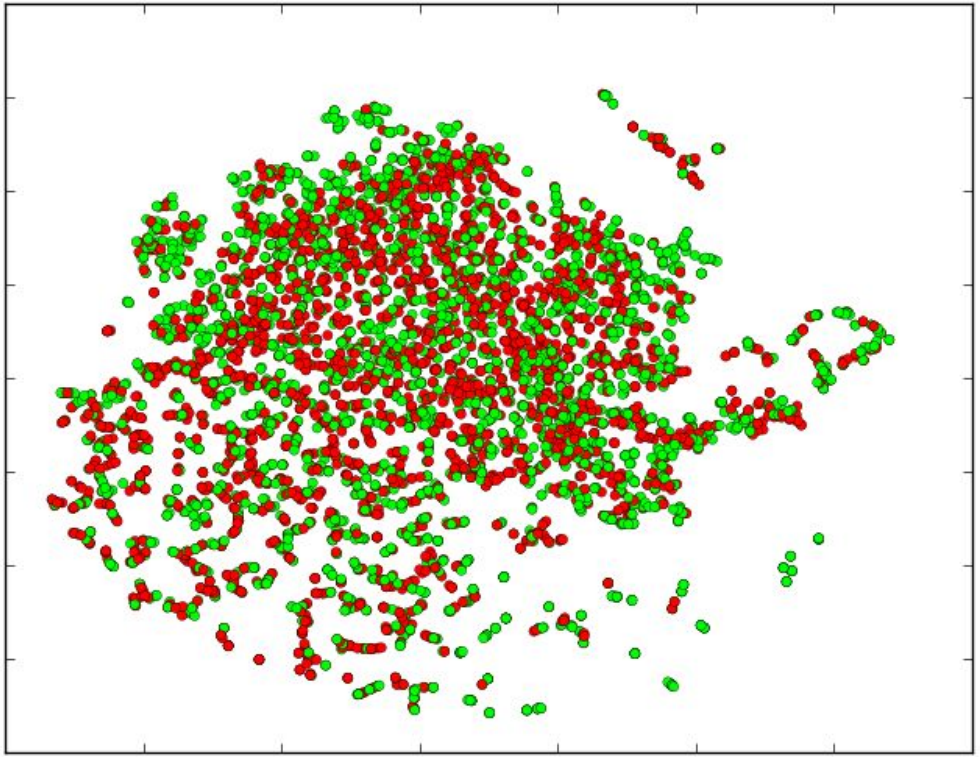
Realizácia v prostredí Microsoft Azure

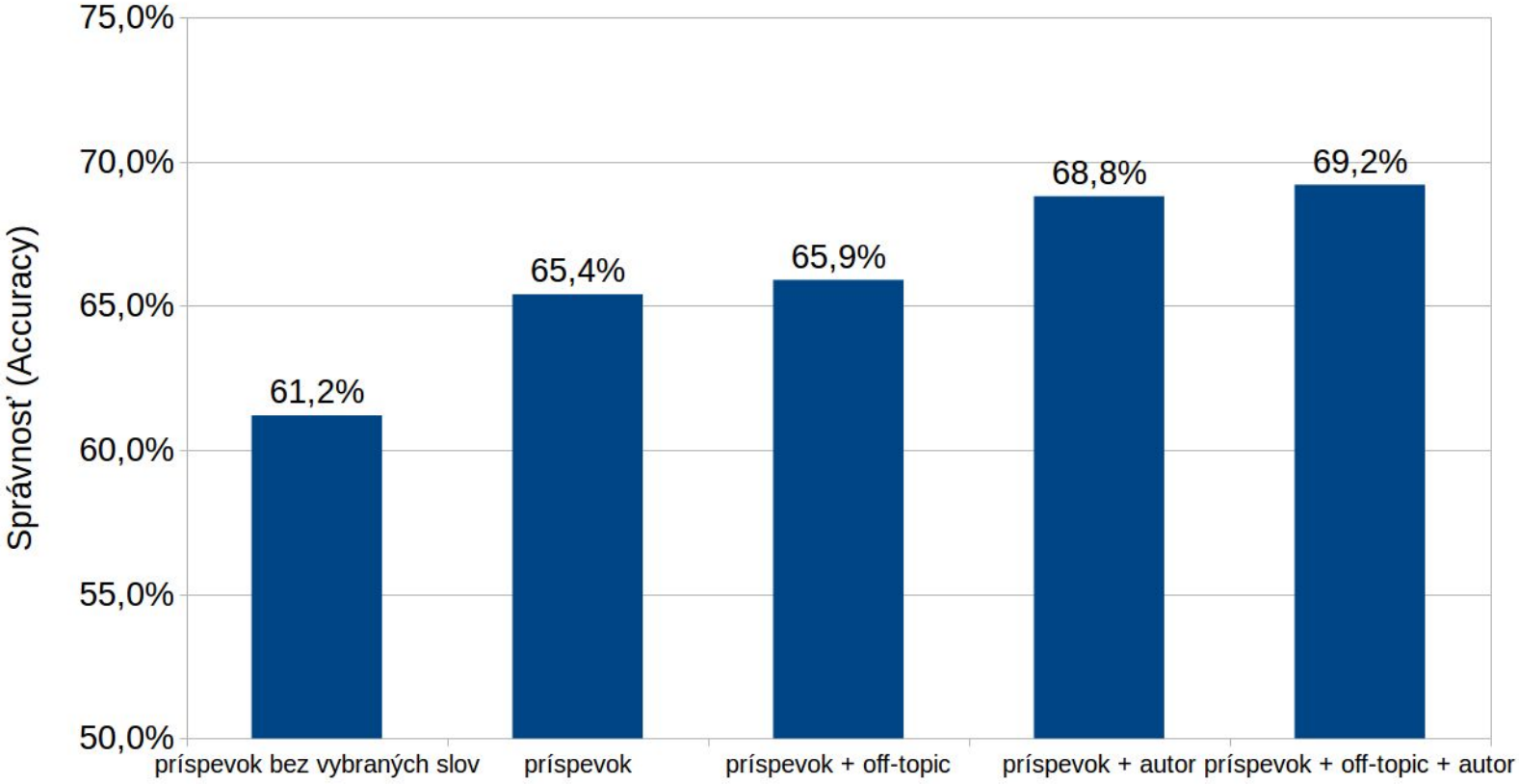


Testovali sme rozličné klasifikátory



Klasifikácia príspevkov je zrejme nelineárny problém



t-SNE vizualizácia

Ďalšie znaky alebo dáta môžu zvýšiť úspešnosť