

Name: Kristina Smirnov

SIGNIFIKANZ-TESTS

Bei Signifikanz-Tests geht darum, herauszufinden, ob zwei Systeme einen signifikanten Unterschied in ihrer Performance (Test Set) aufweisen oder ob die beobachteten Unterschiede rein zufällig sind und nachfolgend welches System bessere Ergebnisse liefert. Dazu muss man zuerst davon ausgehen, dass beide Systeme gleich gut sind (Nullhypothese). Eine sehr beliebte Metrik ist die Accuracy, für die sich der vorhin genannte Sign Test eignet. Eine weitere vorgestellte Metrik ist die Mean Average Precision, die vor allem im Bereich Information Retrieval verwendet wird. Dabei handelt es sich um ein Ranking. Nach dem Testen ist es wichtig, Bias zu vermeiden und ehrliche Ergebnisse liefern, auch wenn diese gegen die eigene These sprechen oder sich als ungünstig erweisen. Man muss sich auch bewusst sein, dass je länger man sucht, desto mehr Signifikanz zu finden ist.