

Obsah

Úvod	7
Použitá symbolika	9
Zpracování dat & digital humanities: Proč dnes?	10
Tabulka plná čísel: matice – nejděsivější z děsivých	15
Korelace a kovariance	20
PCA: Analýza hlavních komponent, vlastní vektory a matice	26
Váhy a rotace	34
Vzdálenosti mezi objekty	43
Jaccardův koeficient	43
Manhattanská vzdálenost	45
Euklidovská vzdálenost	47
Manhattanská vs. euklidovská vzdálenost	48
Kosinová podobnost & nepodobnost	53
Prostorová interpretace euklidovské vzdálenosti a kosinové nepodobnosti	58
Vzdálenosti a číselné vlastnosti	60
Matice párových vzdáleností	66
Vícerozměrné škálování	68
Redukce dimenzí, PCA, SVD, MDS, t-SNE a UMAP	76
Hierarchické shlukování	79
Strojové učení	87
Metoda k nejbližších sousedů (k -NN) a úvodní otázky	90
Motivace k metodám strojového učení upravujícím vstupní vlastnosti	92
SVM	93
Určení důležitosti vstupních vlastností	95



281038423

Lineární oddělovač	97
Určení správnosti klasifikace	99
Férové umístění oddělovače	103
Optimalizační problém	106
SVM v praxi	108
Umělé neuronové sítě (NN)	110
Umělá inteligence (AI), zpracování obrazu a textů	115
Stručný úvod se zápletkou historie	116
Zpracování obrazu	132
Co můžeme s takovými sítěmi dělat?	144
Zpracování textů	150
Sémantické embeddingy	154
Rekurentní neuronové sítě	164
Transformery	173
Co můžeme s takovými sítěmi dělat?	192
Velké jazykové modely, propojení a multimodalita	194
Dnešní doba, data, transfer learning a sítě jako moduly	200
Závěr	207
Bibliografie	209
Rejstřík	212

