

Index

- abstractive, 243
- accuracy, 50, 54
- ad recommendation, 170
- Adam optimizer, 71
- Adobe, 8
- AdWords, 172
- aggregate diversity, 282
- Amazon, 5, 8, 20, 82, 88, 95, 100, 167, 175, 189, 240, 256, 258, 260, 269, 301, 302
- anomaly detection, 212
- area under the ROC curve (AUC), 115, 116, 120, 122, 142, 143
- attention mechanism, 204, 206, 207, 209, 237, 261
 - self-attention, 140, 204, 206, 208
- autoencoder, 128
 - denoising autoencoder, 129
- AutoRec, 120, 129
- autoregression, 178, 212
- bag-of-words representation, 220, 221, 223–225, 227, 229, 235, 249, 251
- balanced error rate (BER), 56, 76
- Bayesian Personalized Ranking (BPR), 114, 134, 135, 137, 142, 199, 275
 - Group BPR, 161
 - Social BPR, 158
 - Visual BPR (VBPR), 254, 269
- Behance, 263
- BERT, 140, 204
 - BERT4Rec, 208
- bigram, 225, 249
- bilinear, 133, 150
- Bing, 289
- bipartite graph, 97, 99, 172
- BookCrossing, 301
- bookmooch.com, 154
- BrightKite, 169
- calibration, 294
- case studies, 100, 183, 193, 210, 234, 241, 247, 257, 286, 300
- Ciao, 156
- Cinematch, 122
- CiteULike, 190
- cold-start, 6, 130, 133, 144, 145, 149–151, 155, 158, 162, 169, 174, 175, 177, 205, 209, 252, 255, 258, 264
- collaborative filtering, 86, 127, 183
- computational social science, 11
- concentration, 14, 273, 274, 277, 278, 304
- concept drift, 188, 211
- conditional independence, 192, 201
- conversational recommendation, 12, 219, 245, 246, 291
- convolutional neural network (CNN), 130, 168, 259, 268
- cosine similarity, 90, 92, 100, 118, 132, 136, 227–229, 233, 250, 278, 280, 287, 292, 304
- Criteo, 140
- crowd-sourcing, 243
- csv, 135
- cumulative gain, 124
- dateutil, 199
- decision tree, 53, 152
- deep learning, 3, 32, 44, 54, 70, 125, 130, 140, 147, 209, 254
- Delicious, 156, 159, 190

- demographic parity, 295, 304
- determinantal point process, 282
- Discounted Cumulative Gain (DCG), 124
 - Normalized Discounted Cumulative Gain (NDCG), 124
- document frequency (df), 228
 - inverse document frequency (idf), 229, 230
- dunnhumby, 166
- EachMovie, 151
- eigenvector, 98, 108
- encoder-decoder, 240
- EndoMondo, 210
- ensembling, 103
- Epinions, 158, 159
- equal opportunity, 296
- Etsy, 8
- Euclidean distance, 126, 127, 132, 133, 164, 197, 198
- explainability, 243, 245
- explicit feedback, 122, 134, 247
- explore/exploit tradeoff, 171
- extractive, 243
- extremification, 14, 274, 275, 277
- f*-test, 74
- Facebook, 274, 278, 288–290
- Factored Item Similarity Models (FISM), 119, 213
- factorization machine, 145, 147, 148, 162, 174, 176, 195, 199, 209, 212, 214, 251
- Factorized Markov Chains (FMC), 195
- Factorized Personalized Markov Chains (FPMC), 193, 195, 196, 199, 213
- fairness, 15, 273, 282, 295, 298–304
- FAISS, 133
- false negative, 55
 - false negative rate (FNR), 56
- false positive, 55, 60
 - false positive rate (FPR), 56
- FastFM, 148, 214
- feature engineering, 32, 219
- feature imputation, 37
- filter bubble, 7, 14, 15, 273–275, 277, 288–290
- Foursquare, 169
- Fraction of Variance Unexplained (FVU), 30
- Gaussian distribution, 27, 30, 46, 73
- Generative Adversarial Network (GAN), 267
- generative model, 231, 252, 267
- Gensim, 232, 250
- Gini coefficient, 276, 277, 286, 304
- Goodreads, 21–23, 38, 46, 102, 135, 148, 175, 275, 301
- Google, 5, 173
 - Google Local, 214
 - Google News, 289, 290
 - Google+, 290
 - Smart Reply, 220, 247
- Gowalla, 169
- gradient descent, 42, 43, 46, 70, 71, 109, 111, 134, 251
- group recommendation, 145, 159
- GroupLens, 141
- herding, 213
- hit rate, 56
- HITS, 98
- Houzz, 262
- hyperparameter, 65, 69, 114, 130, 162, 231, 249, 286, 295
- ImageNet, 255, 258, 266
- Implicit (library), 134, 135
- implicit feedback, 112, 113, 115, 122, 135, 139, 142, 161, 164, 213, 253
- initialization, 111, 136, 142, 149
- inner product, 23, 108, 125, 127, 131, 145, 196, 231, 249, 267, 270, 274, 280
 - maximum inner product, 106, 132, 249
- interactive recommendation, 11, 132, 245, 246
- interpretability, 106, 130, 131, 144, 234, 243, 252
- item bias, 109, 184, 256
- item metadata, 82, 175, 242, 254, 301, 305
- item-to-item compatibility, 118, 142, 272
- item-to-item recommendation, 87, 100–102, 192, 231, 232, 250, 252, 257, 259, 274, 279
- item-to-item similarity, 96, 97, 99, 118, 250, 280
- item2vec, 120, 232, 250, 286
- Jaccard similarity, 88, 89, 91, 92, 95, 97, 101, 133, 142, 175
- json, 25
- KD-tree, 132
- Kiva.org, 298
- Last.FM, 301

324

Index

- latent factor model, 108, 109, 115, 118, 119, 130, 134, 136, 139, 141, 145, 156, 160, 162, 163, 176, 214, 236, 237, 246, 255, 280, 297, 305
- least misery, 160
- LibraryThing, 159
- likelihood function, 28, 51
- line of best fit, 22, 23, 26, 34, 38, 39, 47
- local minima, 43, 111
- locality-sensitive hashing, 133
- location-based social network, 169, 196
- Long Short-Term Memory Model (LSTM), 202, 205, 239, 240
- long-tail, 15, 32, 119, 144, 206, 252, 274
- long-term dynamics, 11, 37, 178, 181–184, 188–190, 201, 204, 205, 210, 212, 214, 256
- Markov chain, 191, 193, 195, 196, 199, 201, 204, 205, 207, 212, 263
- matching problem, 171
- matrix factorization, 84, 106, 107, 125, 157, 158, 193–195, 235
- maximal marginal relevance, 279, 281, 284, 291, 304
- maximum likelihood estimation, 28
- mean absolute error (MAE), 26, 32, 122
- mean reciprocal rank, 123, 142
- mean squared error (MSE), 26, 28–30, 40, 43, 46, 71, 73, 84, 97, 102, 105, 109, 112, 113, 121, 135, 140, 142, 178
 - root mean squared error (RMSE), 26, 105, 122
- mechanical turk, 161
- memory-based personalization, 14, 86, 104, 117, 139, 143, 205, 209
- Microsoft, 8, 232
- Million Song Dataset, 169
- mixture of experts, 265
- ModCloth, 302
- model-based personalization, 13, 19, 86, 96, 104, 106, 112, 117, 139, 142, 205
- MovieLens, 140, 151, 208, 298, 301
- multi-hot encoding, 36
- multilayer perceptron, 3, 53, 126–128, 131, 140, 147, 168, 248
- music recommendation, 168, 232, 286, 292
- n-gram, 224, 249, 251
- neighborhood-based recommendation, 86, 186, 254
- Netflix, 108, 122, 183, 187, 188, 191, 212, 256, 294
 - Netflix Prize, 7, 82, 105, 109, 139, 182, 183, 187, 206, 211
- neural attentive recommendation (NARM), 207
- neural collaborative filtering, 127, 147
- neural factorization machine, 147
- novelty, 139, 168, 291, 292
- one-class recommendation, 113, 237
- one-hot encoding, 33, 36, 39, 41, 62, 76, 146, 148, 150, 165
- online advertising, 152, 170, 173, 298
- online dating, 152, 153, 298
- online learning, 4
- Opaki BM-25, 229
- outfit recommendation, 115, 252, 257, 259–261, 263, 271
- outlier, 19, 30, 40, 77, 121, 178, 186, 189
- overfitting, 62–64, 69, 70, 72, 77, 109, 110, 223
- p*-value, 74
- PageRank, 98
- pairwise loss, 115, 151, 160, 161, 194, 246
- Pearson similarity, 92, 101
- personalized health, 10, 211
- Personalized Ranking Metric Embedding (PRME), 196–198, 200
- Pinterest, 8, 263
- point-of-interest recommendation, 196, 258
- pointwise loss, 115
- PolyLens, 160
- popularity bias, 300
- precision, 58, 60, 76, 281, 297
 - at *K*, 60, 122
 - precision/recall curve, 60, 76
- price dynamics, 53, 144, 162, 163, 165, 175
- pseudoinverse, 23, 25, 26, 42, 47, 66
- Python, 4, 134, 148, 225
- PyTorch, 5
- query refinement, 245
- R^2 coefficient, 29, 46, 62
- random forest, 49, 53
- re-ranking strategies, 281, 284, 291, 302
- recall, 56, 58, 297
 - at *K*, 60, 122
- receiver-operating characteristic (ROC), 61

- recipes, 241
- reciprocal communication, 153
- recurrent neural network, 3, 7, 130, 182, 201, 204, 205, 207, 209, 210, 215, 219, 237, 239
- reddit, 11, 39, 44
- regularization, 64, 65, 67, 70, 72, 76, 109, 116, 118, 119, 131, 136, 142, 143, 149, 156, 164, 214, 227, 235, 249
- residual, 30, 47, 74, 75, 212
- retrieval, 55, 131, 143, 229, 239, 242, 244, 246, 248, 250, 253, 279, 291
- Rossmann, 193
- satisfaction, 106, 139, 160, 161, 244, 284, 288, 291, 292
- scikit-learn, 25
- scipy, 31, 75, 135, 149
- seasonal trends, 167, 178, 180, 181, 191, 215
- semi-supervised learning, 21
- sensitivity, 56
- sentiment analysis, 72, 220, 223, 249, 250
- sequence-to-sequence, 207, 248
- sequential context, 120, 168, 208
- sequential dynamics, 214
- sequential recommendation, 10, 169, 195–197, 200, 204–208, 210, 231
- serendipity, 15, 291, 292
- session-based models, 189, 205
- shared parameter, 157, 164, 236
- short-term dynamics, 178, 181, 189, 205, 210, 212, 214
- side-information, 6, 85, 134, 139, 140, 144, 150, 158, 176, 177, 205, 234
- sigmoid function, 51–53, 112, 116, 127, 157, 158
- similarity-based recommendation, 86, 87, 89, 96, 101, 102, 212
- singular value decomposition, 108, 125
- sklearn, 24, 25, 57, 73, 75, 76, 270
- social connections, 155, 158, 159, 169, 236
- Socially-Aware Personalized Markov Chains (SPMC), 195
- socially aware recommendation, 141, 155, 175, 195
- sparsity, 67, 118–120, 155, 264
- spectrogram, 168
- spline function, 185
- Spotify, 169, 286–288
- Steam, 5
- stemming, 222, 224
- stopwords, 222, 226, 227
- Support Vector Machine (SVM), 49, 53
- Surprise (library), 134
- t-SNE, 233, 270
- Taste Profile Dataset, 169
- temporal dynamics, 39, 140, 177, 180, 183, 187–191, 211, 213, 256, 263
- Tensorflow, 4, 70, 134, 136, 137, 141, 199, 266, 299
- term frequency (tf), 228
- test set, 63, 66, 68, 69, 73, 105, 109, 123, 135, 149, 174, 176, 270
- text generation, 202, 209, 237, 238, 240, 241, 243, 247
- tf-idf, 229, 249, 281
- Theano, 5
- tokenization, 233
- topic model, 235–237, 262
- Tradesy, 256
- training set, 64, 66, 69, 70, 73, 77, 102, 111, 116, 123, 159, 174, 176, 238, 261, 267, 276
- Transformer, 140, 206, 208
- transition probability, 98
- translation-based recommendation, 197, 200
- trigram, 225
- true negative, 55, 60
 - true negative rate (TNR), 56, 57
- true positive, 55, 60
 - true positive rate (TPR), 56, 57
- trust, 13, 155, 156, 158, 244
- tsv, 135
- Twitter, 289
- underfitting, 62, 64, 69
- underrepresented group, 15, 273, 305
- unexpectedness, 291–293
- unfairness, 296, 297, 301, 303
 - absolute unfairness, 297, 299, 300, 303, 305
 - value unfairness, 296, 305
- unigram, 225, 249
- unix time, 82, 200
- unsupervised learning, 21, 267
- user bias, 185
- user interface, 11, 122, 139, 141, 160, 181, 183, 191
- user manifold, 3, 6, 12
- user-free personalization, 117, 120, 129, 133, 205, 231
- user-to-item relationship, 100, 145, 173, 287

326

utility gap, 300

validation set, 66, 69, 70, 73, 77, 223, 249
visual attributes, 252, 262
visual compatibility, 257, 258, 262, 263
visual dynamics, 262, 263
visual signals, 252, 254, 256
visual similarity, 257, 258

Index

visually aware recommendation, 7, 8, 254, 257

wide & deep model, 147
word2vec, 120, 230, 232

Yahoo, 289
Yelp, 20, 74, 140, 240, 244
YouTube, 169, 254, 274, 278, 282, 286