

Index

A

abstractive summarization 31, 54, 166, 180, 185, 208, 251
 advertising 186-190, 197, 201-204
 archaeology 76, 77, 82, 92, 95, 96
 aspect-driven summarization 282, 290, 293
 Association for Computational Linguistics (ACL) 24-27, 51-53, 72-74, 93, 94, 132-137, 140, 157-162, 181-183, 202-204, 228-231, 251-255, 273-276, 290-296, 306-308
 automatic summarization 2, 25, 53, 68, 75, 97, 112, 129, 133-136, 140, 141, 145, 157-160, 163, 183, 185, 188, 206, 207, 226-230, 239, 250, 251, 259, 273, 292, 296-304
 automatic summary evaluation 98, 102, 109, 110, 123, 128, 132, 156, 270, 271
 automatic text summarization 2, 25, 73, 93, 98, 101, 102, 129, 133, 136, 137, 159, 160, 164, 166, 170, 180-188, 202, 204, 227-230, 233, 258, 259, 273-277, 291-295, 305

B

bag-of-words approach 3, 5, 11-13, 20, 23, 46, 209-212, 234, 237, 302
 BLEU evaluation metric 109-111, 115, 119, 123, 127, 128, 131, 135, 138, 270, 274, 289
 burstiness score 152, 153
 burst pattern 147, 149, 236, 240-243, 251-256
 bursty subjects 236

C

clustering methods 20, 193
 contextual advertizing (CA) 27, 28, 132-136, 162, 183-192, 197, 200-204, 229, 309
 convex polytopes 55, 71, 75
 cross-language document summarization 53, 142, 255, 261, 275, 277, 293-296

D

digital preservation 97
 document clusters 7-10, 17, 18, 22, 31-34, 56, 57, 92, 95, 102, 126, 145, 148-152, 168-171, 178, 209-211, 217-220, 227, 236, 260, 261, 265-268, 272, 279, 286, 290
 Document Understanding Conference (DUC) 2-4, 7-10, 13, 16-25, 29, 30, 34, 36, 41, 42, 45, 55, 67-75, 89, 93-95, 101, 102, 110, 113, 116, 118, 125-128, 132, 133, 155, 156, 171, 172, 181, 183, 202, 208, 210, 213, 226-231, 235, 238, 239, 256, 268-271, 274, 276, 279, 280, 292-296, 304-306, 309
 domain-specific summarization 32, 54

E

extractive multi-document summarization 25, 211
 extractive summarization 3, 30-32, 54-58, 71-75, 166, 170, 180, 185, 207, 208, 228, 230, 233, 234, 239, 250, 252, 255, 265, 274, 299, 307

F

feature selection 145, 202, 205

G

generic summarization 29-34, 54, 137, 169, 175, 234, 255, 260

H

hyperplanes 58-61, 65, 67, 71

I

indicative summaries 102, 137, 259, 260

information extraction (IE) 76-78, 92, 94, 97, 141, 143, 154-157, 162, 163, 181, 187, 202, 210, 252, 279, 282, 291, 296, 300, 303, 306-309
 information retrieval 24-27, 30, 32, 52, 53, 58, 72-74, 79, 93, 95, 101, 103, 106, 132-137, 141-147, 154, 158-162, 169, 173, 181-184, 187, 188, 195, 201-205, 226-231, 237, 243, 252-255, 266, 273-276, 291-296, 304-307
 informative summaries 118, 137, 260
 interactive summarization 97, 249
 International Workshops on Semantic Evaluation (SemEval) 140

K

keyword extraction 79, 80, 97, 252

L

language-independent sentence features 255
 language-independent summarization 34, 54, 224, 269, 297, 299
 language-specific sentence features 255
 latent semantic analysis (LSA) 24, 28, 91-94, 182, 203, 204, 229, 236, 273, 278-284, 287-295
 parallel corpus 235, 278, 279, 286-290, 293
 statistical machine translation 108, 141, 189, 278, 280, 286
 linear optimization 75, 265
 linear programming (LP) 55-57, 63, 67, 70-75, 210, 235, 251, 295
 linguistics 5, 17-20, 25-28, 51-53, 66, 73, 74, 94, 97, 103, 132-137, 145, 157-162, 181-185, 202, 204, 213, 226-231, 251-255, 273-276, 292-296, 306, 308

M

microblogging 147, 160, 163
 mobile applications (Apps) 208, 223, 231
 monolingual summarization 54, 297, 299
 multi-document summarization (MDS) 1-4, 24-39, 49-57, 67, 68, 72-74, 95, 135, 143, 146, 159, 168, 173, 181-184, 188, 202, 210, 211, 225-234, 239, 249-255, 260, 265-269, 275-277, 292, 296, 307, 308
 multi-document summarization (MDS), multilingual 278
 multilingual document clustering 277
 multilingual documents 258, 277

multilingual news summarization 206, 227, 273
 Multilingual Sentence Extractor (MUSE) 35-40, 46, 50, 51, 54, 211, 234, 237, 238, 249, 250, 256, 265
 multilingual summarization 52, 73, 182, 228, 231, 234, 258-261, 265-268, 272, 274, 277-280, 288-299, 308
 multilingual summarization evaluation 279, 288, 293
 multilinguism 309

N

National Institute of Standards and Technology (NIST) 2-4, 22-25, 30, 51-54, 68, 72, 93, 94, 109, 110, 116, 125-135, 155, 181, 227, 228, 231, 251, 254-256, 271, 277, 291-296, 304, 305, 309
 natural language processing methods 17, 26
 natural language processing (NLP) 4, 5, 17-23, 26-28, 52, 53, 71-75, 89, 94-97, 103, 122, 128, 132-137, 140-145, 154, 157, 160, 162, 166, 170, 180-185, 202, 226-229, 251-255, 271-276, 281, 293, 297, 306-309
 news summarization 52, 74, 142, 146, 165, 173, 176, 206-213, 225-235, 239, 254, 273, 291, 292, 296
 NewSum summarization application 206-208, 213-225, 231
 N-gram graph framework 4, 21, 37, 66, 68, 102, 108-111, 119, 131-134, 137, 138, 153, 171, 204, 206, 209-220, 225-228, 231, 236, 252, 254, 270-274, 290, 303, 307

O

opinion summarization 73, 140, 144, 154, 155, 164, 165, 168-180, 185, 208

P

parallel corpus 235, 278, 279, 286-290, 293
 polyhedrons 61, 62, 75
 polytopes 55-58, 61-65, 71-75
 probabilistic topic modeling methods 12, 23, 26

Q

query-based summarization 54, 72, 234, 256

R

recall 12, 15, 30, 34-37, 45-50, 65, 68-71, 87-90, 99, 103-106, 110, 113, 114, 119, 122, 123, 131, 137, 152, 155, 156, 179, 180, 191-196, 200, 217, 238, 245, 269-271, 284
Recall-Oriented Understudy for Gisting Evaluation (ROUGE) 4, 10, 13, 16, 21, 22, 25, 27, 34, 37, 45-51, 54, 68, 71, 73, 89, 94, 110, 111, 116, 119, 123, 126-128, 131, 133, 138, 155, 156, 160, 171, 172, 178, 182, 185, 210, 235, 238, 239, 242, 256, 269-273, 284, 285, 289-291, 296, 304-309
redundancy 3, 6, 13, 26, 34, 35, 47, 57, 78, 144, 168-171, 185, 206-214, 218-220, 230, 282, 290, 300-302, 308
reference summaries 4, 21, 37, 68, 83, 84, 87-89, 101, 102, 111-118, 125, 129, 155, 180, 185, 270-272, 285, 304-306
ROUGE-1 scores 29, 30, 34, 46-50, 70, 238, 239, 271
ROUGE-2 scores 29, 35, 45, 46, 49, 50, 70, 129, 155, 156, 271, 289, 302-305

S

sentence extraction 23, 70, 73-75, 81, 97-100, 135, 160, 204, 227, 229, 239, 275, 276, 293, 295, 299, 307
sentence ranking 1-17, 20-26, 32, 99, 234, 235, 263, 265
sentence relations (SentRel) multi-document summarization algorithm 29-32, 37-39, 44, 48-51
single document summarization 33, 71, 110, 168, 260, 277
singular value decomposition methods 11, 23, 26
social computing 163
social events 150, 163
social media 140-147, 150-163, 206, 208
social media, semantic analysis in 141, 157, 159, 162, 163
social media summarization 140-147, 156, 157, 163
social media text 143-146, 163
social networking service 163
stemming 5, 50, 57, 60, 71, 74, 75, 148, 198, 209, 215, 299
stop-words 5, 6, 13, 37, 50, 57, 66, 71, 75, 80, 123, 137, 166, 167, 173, 178, 185, 190, 198, 209, 262, 268, 285, 287, 299
SumBasic 27, 176, 185

summarization algorithms, supervised 54
summarization algorithms, unsupervised 54
summary evaluation 22, 24, 35-37, 94, 98-106, 109, 110, 118, 119, 123-128, 131-133, 136, 137, 156, 214, 229-234, 237, 268-272, 285-290
summary generation 3, 4, 8, 12, 13, 21, 23, 26, 31, 171, 259, 260, 302
supervised learning 6, 166, 167, 185
symmetric directed graph 38-40, 54

T

temporal text mining 232, 233, 236, 240, 243, 252, 253, 256
Text Analysis Conferences (TAC) 2-7, 21-25, 29, 30, 34, 35, 40, 46-54, 73, 74, 89, 93, 94, 128-131, 135, 140, 155, 156, 159, 168, 171, 172, 181-185, 208, 213, 226-228, 231, 254-256, 268, 269, 272, 273, 277-296, 301, 305-309
Text Retrieval Conferences (TREC) 101, 106, 125, 126, 136, 140, 146, 155, 156, 228, 236, 251, 253
text summarization 2, 3, 11, 24-27, 30, 51, 55, 73, 74, 79, 81, 89-94, 97, 98, 101-105, 129, 132-137, 159, 160, 164-172, 175-205, 227-230, 233, 239, 250, 251, 254, 258-261, 265-269, 272-277, 291-298, 305-309
tokenization 137
Twitter 141-162, 170, 182, 250, 253

U

update summarization 27, 140, 146, 156-159, 163, 227, 235, 243, 253, 292
user comments 145, 164, 168, 170
user-focused summary 137

V

vector models 137, 212
VERT evaluation metric 119, 123-131

W

word clustering 5, 8-10, 20, 23
word clustering methods 26
word frequency methods 26
word tokens 3-7, 83, 91, 137, 152, 216, 217, 236, 264, 299, 302, 303