Index

A

a-priori-based algorithm 98
a-priori-based approach 83
access control 261
ActiveXML 259
active caching 105, 115, 119
active server page (ASP) 140, 266
active XML 115
actual mining 92
adaptive resonance theory (ART) 20
adhesion constraints 90
Akamai 129, 143
Akamai EdgeSuite 130
ALIAS 207
anomalies detection 191
Apache 108, 131, 248
application fields 165, 168
application program 131
application programming interface (API) 110, 143, 275
application server 128
application tuning 192
artificial intelligence (AI) 18, 256
artificial neural network (ANN) 4, 18, 21
ART family of networks 21
aspect-oriented programming (AOP) 273
association-rule mining 98
association rule 79
associative clustering neural network (ACNN) 21
automatic answering 229
automatic mail organization 230
autonomic 277
AXML 260, 261

B

bag-of-words 223, 224
BASIL 254
Bayesian neural network (BNN) 22
BEA 145
BEA WebLogic 128
BPEL4WS 257
browse 127
business-to-business (B2B) 245, 267

C

cache-control 112, 134
cache-worthiness 149, 150, 152
CachePortal 143
cache coherency 133
cache server 138
caching 105, 111-114, 125

Copyright © 2007, Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.
Index

caching dynamic content 115
caching middleware 115
caching sScores 152
candidate item set 88
CASE tool 180, 183, 185
categorization 226
centralized management 273
choreography 256
classes 229
classical algorithms 83
classic algorithms 98
classification 80, 222
client-based solution 233
client polling 135
client side 114
clustered server 117
clustered Web 117
clustering 2, 35, 74, 80, 117
clustering in quest (CLIQUE) 5
clustering interpretation 47
clustering large applications (CLARA) 10
clustering with snippets 26
cluster features (CFs) 11
coarse-grained dynamic-content caching 112
coherence 134, 138
combined methods 17
common data model (CDM) 200
common gateway interface (CGI) 105
common object request broker (CORBA) 244
communication 138
communication-channel congestion control 192
community 43
compactness 35
comparison of prediction schemes 171
completed delta 94
complex-structured XML documents 91
complex caching 117
complex document 84
complex XML document 81
composite system 270, 279, 283
computational linguistics 222
computational overhead 137
conceptual logs 185, 187
confidence 82, 83
consolidated delta 92, 95, 98
constraints 90
content analysis 80
content caching 111
content change 99
content delivery 129
content delivery network 49
content delta 83
content dynamism 114
content fragment 116
content middleware 105-110
content personalization 180
content unit 183, 194
context selection 85
controller 277, 282
cophenetic correlation coefficient (CPCC) 47
copyright protection 132
correspondence analysis (CO-AN) 48
correspondence assertion 214

d
data-integration 201
database accelerator 129
database administrator (DBA) 201
database approach 253
database server 129
database system (DBS) 200
data clustering 1
data exchange 261
data mining 220, 222
delta 92
digital search tree(s) 163
directed graph 13
directory based 42
distance-space method 7
distance measure 5
distribution 261
distribution network 129
document-constructor model 214
document index graph (DIG) 25
document representation 62, 67
DTD 59, 65
DUP 142
Dynamai 149
dynamic 140

Copyright © 2007, Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.
dynamic-content caching 140
dynamic-content delivery 119
dynamic-content middleware 119
dynamic association rule 98
dynamic content 105, 111, 117, 118, 125
dynamic data 99
dynamic mining 99
dynamic Web content 105, 106, 109, 112
dynamic Web content middleware 111
dynamic Web content technology 119
dynamic Web page 125, 141
dynamic XML document 83, 92, 98

E
e-mail 220, 222, 226, 228
e-mail clustering 228
e-mail header 224
e-mail management 219
e-mail mining 49, 220, 226
EACH 207
early server-side practices 109
Edge Side 130
Edge Side Includes (ESI) 143
edit distance 63
edit operation 63
EFLWOR expression 205
enterprise application integration (EAI) 244
entity-relationship (E/R) 183
EQUAL 207
evaluation criteria 271
expectation-maximization (EM) algorithm 45
extended work-flow approach 257
extracting association rules 85, 87

F
feature 67, 70
feature extraction 2
feature selection 2, 224
feature vector 4
federated database 201
feedback controller 278
fine-grained caching 114
fine-grained proxy caching 113
first in, first out (FIFO) 132
folders 230
fragmentation 105, 109, 111, 114
fragments 116
frequently changing structures 97
freshness 105, 117, 118
FULL 208
functionality based 42
functional properties 250
fuzzy cluster analysis 10

G
generic association rule 92, 99
global-as-view (GAV) 201
granularity 70, 75
granularity level 59
graph 40, 42, 44
graph clustering 12
GreedyDual 133
growing cell structure (GCS) 20

H
heterogeneous data source 204
hidden knowledge 79
HIDE 208
hierarchical agglomerative algorithm 10
hierarchical clustering 3
historical structural deltas 93
historic change 99
historic version 83, 92
hypergraph 191
hypertext 183
hypertext markup language (HTML) 105, 222, 266
hypertext transfer protocol (HTTP) 245
hypothesis (Ho) 47

I
IBM WebSphere Application Server 128
IBM WebSphere Cache Manager 128
IBM WebSphere Edge Server 130, 149
If-Modified-Since (IMS) 134
incremental clustering 4
information-retrieval 253
information-theoretic method 159
input signature 251
integrating databases 200

Copyright © 2007, Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.
integrating heterogeneous data sources 199
integration strategy 211
interesting information 99
interesting knowledge 80, 97
interesting relationships 80, 82
interstructure mining 80
intrastructure mining 80
item 82, 91

J
Java 2 Enterprise Edition (J2EE) 244
Java Caching System (JCS) 131
Java Edge Side Includes (JESI) 144
Java server pages (JSP) 266
Java specification requests (JSR) 131
JSP Cache Tag Library in BEA WebLogic 145
JSR-107 131

K
k-means algorithm 228
KDD 181-188
KEY 208
knowledge 92
knowledge-based approach 254
knowledge discovery 80
k clusters 228

L
large item set 82,84
Large Scale Distributed Information Systems lab (LSDIS) 274
lease 139
least likely to be used (LLU) 132
least recently used (LRU) 132
Lempel-Ziv-78 scheme 165
level constraint 90
link-based approach 44
link-based clustering 25,42
link-based similarity 73
links 59,62,73
local-as-view (LAV) 201
local area network (LAN) 245
location processor (LP) 211
LRUMIN 133

M
machine learning 220, 223, 226, 233
MAFIA 5
mail retrieval agent (MRA) 221
mail summarization 231
mail transfer agent (MTA) 221
mail user agent (MUA) 221
management aspects 273,279
management framework 270, 273, 283
mapping 64
Markov chain 161
Markov predictor 163
MDBS context 204
mediator 201
Microsoft .NET 244
middleware 244
minimum-cost edit script 94
minimum confidence 82,86
minimum support 82,86
mining association rule 99
mining Web applications 179
misleading statistics 131
mobility tracking 160
model-based approach 44
model-based scheme 45
multicasting of dynamic content 116
multidatabase 201
multidatabase language (MDL) 200
MXQuery 200,204

N
n-gram 232
naïve bayes classifier 226
NAME(n) function 208
names 62
natural language processing 222
network overhead 137
non-a-priori based approach 83, 89
nonfunctional properties 250
NULL 208

O
object request broker (ORB) 244
ontology 180
optional sources 214
Oracle 9iAS Web Cache 128, 142
Oracle Application Server 128
Oracle Web Cache 149
orchestration 256
ordered model 94
Organization for the Advancement of
Structured Information Standards
(OASIS) 268
origin server 138
output signature 251
OWL-S 259

P
page annotation 180, 182
partitioning around medoids (PAM) 9
path 59, 67, 70
patterns 80, 99, 181, 188, 195
patterns statistically 189
performance 107, 109
performance accelerating 117
Persistence software 146
piggyback client validation (PCV) 136
piggyback server invalidation (PSI) 136
polling-every-time 135
polymorphism 114, 119
portal 125, 149
position constraint 90
postprocessing 92
postprocessing association rule 85
postprocessing phase 88
pre-processing 181
prediction-by-partial-match scheme 164
prediction models 161
preprocessing 223
preprocessing data 85
preprocessing phase 85
preprocessing the document 84
privacy 132
Privacy Preferences Project (P3P) 283
probabilistic latent semantic analysis
(PLSA) 182
probabilistic suffix tree scheme 167
proxy caching of dynamic Web content 112
proxy log-file 39
proxy server 127
pull 138
pull-based coherency mechanism 135
push 138

Q
quality of data (QoD) 105, 117
quality of service (QoS) 267, 271, 282
queried source 214
query optimization 261
query postprocessor (QPp) 211
query processor (QP) 211
query user interface (QUI) 211

R
Rand statistic (R) 47
recency 132, 151
records 84
recovery 272
reduction of dimensionality 5
reflective and adaptable Web service
( RAWS) 274
relational association-rule algorithm 88
relational database 80
relational mining techniques 81
remote procedure call (RPC) 244
representation 223
resilience 138
resource provisioning 272

S
scalability 138
scripting language 108
secure sockets layer (SSL) 275
security 132, 261, 267, 272, 275
self-configuring 277
self-healing 277
self-optimizing 277
self-organizing map 19
self-organizing neural network 19
self-organizing tree 20
self-organizing tree algorithm (SOTA) 20
self-protecting 277
semantics 213
semantic heterogeneity 204
semistructured data 79
semisupervised learning 229

Copyright © 2007, Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.
separation 35
sequential minimal optimization (SMO) 228
server-based solution 233
server-side acceleration 117
server accelerator 128
server side 110
service-level agreement (SLA) 267
service-oriented architecture (SOA) 245, 275
service-oriented computing (SOC) 243, 267, 283
service-oriented middleware (SoM) 274
service discovery 252
service matchmaking 252
session 40, 44
similarity 5
similarity-based algorithm 45
similarity-based approach 44
similarity coefficient 6
similarity measure 2
singular value decomposition (SVD) 5
site-based LRU 133
size-adjusted and popularity-aware LRU (LRU-SP) 133
size-adjusted LRU (SLRU) 133
size strategy 133
SLA 271, 276
smart-home 167
SOAP 247
space overheads 137
spam 233
spam filtering 233
spam mail 233
staleness 117
static XML documents 83, 91
stemming algorithms 223
structural change 99
structural delta 83
structure clarification 80
structure of clusters 25
subspace clustering 5, 12
suffix-tree-based prediction scheme 172
suffix-tree clustering (STC) 24
support 82, 83
support-vector machines 227
syntax 213

T
tag 58, 72
tagged element 58
technical approach 233
template 109, 181, 187, 188, 194
term-based clustering 26
term-document matrix 5
text 67
text-based approach 44
text-based clustering approach 42
text mining 222
thread summarization 231
time-series-based approach 72
time-to-live (TTL) 134
time-to-refresh (TTR) 135
traffic analyzers 181
transaction 82, 84
transductive support-vector machine 229
transparent proxy 129
tree 60
tree-based approach 62
tree-structured item 91
tree edit distance 60, 65
two-way joining algorithm 11

U
undirected graph 13
universal description, discovery, and integration (UDDI) 243, 248, 268
unordered tree model 94
unordered XML tree 95
usage-mining preprocessing 187

V
validation 37, 47
variable-length predictors 171
vector 60
vector-based approach 67
vector-based similarity measure 68
vector-space model 223
volume lease 139

W
W3C (World Wide Web Consortium) 84, 269
Index 303

Web 187
Web-log 180, 182
Web-service-based data management 261
Web-service composition 256
Web-service composition language 257
Web-service discovery and match-making 250
Web-service infrastructure 245
Web-service management frameworks 270
Web-service standards 247
WebML 182, 185
Web application development 182
Web content mining (WCM) 22
Web data source 37
Web document 35, 40
Web document clustering 35
Web knowledge discovery area 79
Web modeling language (WebML) 180
Web page 4
Web personalization systems 48
Web portal 149
Web prefetching 49
Web proxy log preprocessing 41
Web reality 200
Web search engine 49
Web server-side log file 38
Web services 244, 259, 266, 268
Web services management 266
Web service definition language (WSDL) 243
Web service management network (WSMN) 274
Web service offering infrastructure (WSOI) 275
Web service offering language (WSOL) 275
Web source 35
Web structure mining (WSM) 22
Web usage mining 22, 180
Web user 35, 38
Web user clustering 35
well-formed tree 90
word-based and flat clustering 26
word-based and hierarchical clustering 26
work-flow approach 255
World Wide Web 267
WS-CDL 258

WS-security 248
WSDL 248
WS discovery 262

X
Xalan 87
XML 186
XML-enabled association rules 91
XML-specific approaches 65, 70
XML (extensible markup language) 79, 105, 200, 243
XML association rules 82
XML data 80
XML document 43, 58, 79, 204
XML mining 80
XML warehouse 94
XQuery 204
XQuery implementation 87
XQuery language 84
XQuery subquery 214