Index

ACRE-learnable, 49, 295
action, 20, 21, 27, 295
additive gap, *see* approximate optimality, additive gap
additive optimality, *see* approximate optimality, additive optimality
adversarial learning, 18, 56, 104, 295
allergy attack, 53–54
anomaly detection, 4, 8, 18, 26, 28, 37, 41, 52, 69, 70, 134–138, 144, 200, 243, 295
  anagram, 47
  Antidote, 9, 52, 136, 137, 147–149, 151, 156, 157, 158, 161, 162, 163, 244, 250
approximate optimality
  additive gap, 207, 295
  additive optimality, 206–210, 212, 217, 225, 289, 295
  multiplicative gap, 207, 210, 213, 220, 295
  multiplicative optimality, 206–212, 217, 227–230, 294, 295
attack categories, 34–47, 49–56, 58, 61, 62, 64, 65, 69, 71, 103, 105, 106, 112, 113, 115–120, 130, 131, 134, 135, 144, 162, 199, 242–244, 246, 298, 304
  availability, 34–36, 51, 54, 62, 120, 242
  causative, 34–36, 38, 43, 46, 49–52, 54–56, 58, 61, 62, 65, 69, 71, 103, 105, 112, 115, 120, 131, 134, 242, 243, 246, 304
  causative availability, 51, 53, 54, 105, 106, 112, 113, 116, 118–120, 298
  causative integrity, 51, 52, 119, 135, 144, 162
  exploratory, 34–36, 38–40, 42, 43, 45–47, 50–52, 61, 62, 64, 199, 242, 246
  exploratory availability, 44
  exploratory integrity, 40, 54, 112
  indiscriminate, 34–36, 40–42, 45, 51, 53, 54, 113, 116, 117, 130, 242
  indiscriminate availability, 115
  indiscriminate causative availability, 131
  integrity, 34–36, 40, 46, 47, 51, 61, 62, 71, 103, 105, 134, 242, 243
  privacy, 34–38, 61, 62, 65, 242, 244
  repeated, 71
  targeted, 34–36, 40–42, 45, 46, 51, 53, 54, 71, 113, 116, 118, 134, 242
  targeted availability, 115
  targeted integrity, 115
attacker, 26, 36, 296

backbone network, *see* computer networking, backbone network
beta function, 109, 296
binary classification, 26, 28
blind spot, 18, 199, 296
bootstrap retraining, 75, 96, 103, 104

chaff, 134, 136, 140–143, 296
classification, 23, 26, 296
classifier, 26, 296
computer networking
  backbone network, 134, 136, 137, 138, 150
  origin-destination flow, 137–140, 143, 150, 151, 153, 159, 162
  point-of-presence node, 137, 138, 140, 141, 151
convexity, 176, 207, 232, 233
  convex body, *see* convexity, convex set
  convex combination, 76, 108, 257, 297, 297
  convex function, 177, 179–182, 184, 186–189, 192, 194, 210–212, 216, 221, 231, 259, 288, 290, 293, 297, 297
  convex hull, 213, 214, 227, 227, 228, 285, 288, 290, 291, 293, 297
  convex optimization, 94, 147, 177, 184, 195, 211, 232, 297
  convex set, 200, 201, 210, 211, 212, 215, 218, 221, 222, 222–225, 231, 232, 257, 259, 297, 297, 302
  convex-inducing classifier, 11, 18, 200–202, 206, 209–211, 218, 225, 231, 232, 244, 245, 289, 297
correlated outlier attack, 54
cost function, 32, 297
covering, 290, 292–294, 297
  covering number, 205, 230, 260–265, 293, 297

https://doi.org/10.1017/9781107338548.016 Published online by Cambridge University Press

data, 20, 297
data collection, 23, 45, 297
data point, 23, 26, 27, 297
data sanitization, 18, 297
database, 171, 175, 176, 178–180, 183, 186, 197
dataset, 23, 25, 26, 298
decision space, 20
deep neural network, 44, 48, 298
defender, 26, 36, 298
degree of security, 13
denial-of-service attack, 9, 18, 44, 52, 134, 138, 140–144, 146, 151–153, 163, 243, 298, 303
dictionary attack, 112, 113, 298
differential privacy, *see* privacy, differential privacy
dispersion, 136, 145, 145, 147, 148, 298
distributional robustness, 136, 298
distributions
  beta distribution, 109, 110, 296
  Erlang distribution, 182, 183, 188, 298
  Gaussian distribution, 25, 136, 159, 160, 235, 299
DNN, 44, 48

explanatory variable, 23, 298

false negative, 27, 31, 32, 34, 36, 51, 114, 115, 119, 134, 156, 157, 299
false-negative rate, 27, 52, 96, 105, 120, 134, 137, 151, 157, 162, 299
false positive, 27, 32, 34, 35, 71, 114–117, 156, 157, 299
false-positive rate, 27, 28, 102, 105, 109, 120, 131, 134, 151, 157, 162, 243, 299
feature, 22, 30, 31, 38, 46, 47, 52, 63, 113, 212, 213, 219, 234, 246, 299
feature deletion attack, 41, 46, 299
feature selection, 24, 30, 31, 41, 46–47, 246, 247, 299
feature selection map, 21, 24, 46, 72, 236, 247, 299

good word attack, 41, 42, 299

Hilbert space, 102, 178, 178, 188, 300, 301
  reproducing kernel Hilbert space, 177–179, 185, 186, 188, 189, 197, 198, 233, 300
hypothesis, 21, 22, 24–27, 300
hypothesis space, 24–26, 47–48, 300

index set, 256
indicator function, 256, 300
inductive bias, 22, 24, 300
inductive learning, 22, 300
input space, 23, 300
intrusion detection system, 31, 40, 41, 45, 47, 53, 201, 300
intrusion instance, 26, 300
intrusion prevention system, 31, 301
iterated game, 58, 71, 301

kernel function, 39, 64, 69, 102, 177–180, 182, 186, 187, 189, 191, 192, 197, 198, 232, 233, 301
  Gaussian, *see* kernel function, RBF
  linear, 194, 196
  RBF, 64, 175, 182, 197, 301
  translation-invariant, 178, 187–189, 191, 193, 194, 198

label, 23, 25–28, 33, 55, 58, 59, 63, 106, 107, 109, 110, 110, 114, 115, 120, 130*t*, 201, 235, 301
labeled dataset, 23, 301
Laplace distribution, 173, 182
Laplace noise, 173, 179, 186, 188, 198, 244
latent Dirichlet allocation, 53
learner, 24, 301
learning procedure, 25, 47–48, 301
loss function, 21, 21, 25–27, 40, 50, 58, 59, 59, 111, 112, 174, 177, 179, 180, 182–184, 186, 187, 192, 301
  0-1 loss, 59, 177
  hinge loss, 177, 184, 185, 187, 192–194, 196–198

machine learning, 20, 302
malfeasance detection, 4, 302
measurement, 22, 302
measurement map, 23, 302
median absolute deviation, 136, 147, 149, 302
membership query, 203, 231, 245, 302
mimicry attack, 41, 47, 54, 201, 302
minimal adversarial cost, 204, 302
Minkowski metric, 210, 302
multiplicative gap, *see* approximate optimality, multiplicative gap
multiplicative optimality, *see* approximate optimality, multiplicative optimality

near-isotropic, 222, 223, 224, 226, 302
near-optimal evasion problem, 11, 48, 200, 205, 302
negative class, 26, 210, 211, 225, 231, 232, 302
norm, 257, 302
  Hilbert space norm, 173
  L^1 norm, 175, 179, 187, 204
  L^∞ norm, 171, 185
  L^p norm, 203, 226
  Mahalanobis, 102
normal instance, 26, 302

obfuscation, 8, 11, 40–42, 46, 112, 118, 302
Ockham's Razor, 22, 303
OD flow, *see* computer networking, origin-destination flow
OD flow volume anomaly, 138, 303

https://doi.org/10.1017/9781107338548.016 Published online by Cambridge University Press

one-shot game, 58, 303
online learning, 25, 59–61, 122, 235, 249, 252, 303
  expert aggregation, 56, 60, 133, 249, 252
  experts, 56, 58, 59, 59, 60, 249, 250, 252, 298
overfitting, 27, 303

PCA, *see* principal component analysis
  PCA-Grid, 136, 145, 146, 146–149, 150, 154, 156
PCA evasion problem, 142, 143, 303
  add-more-if-bigger poisoning attack, 141, 155, 163
  boiling frog attack, 144, 151, 159, 160, 160, 161, 162, 163, 296
  globally informed attack, 141, 142, 146, 156, 163
  single training period attack, 143, 151, 153, 155, 156, 157
performance measure, 26, 303
polymorphic blending attack, 41, 303
PoP node, *see* computer networking, point-of-presence node
positive class, 26, 203, 210, 211, 225, 231, 288–290, 293, 303
positive homogeneous function, 214, 303
prediction, 25, 303
principal component analysis, 9, 18, 37, 52, 134–145, 146, 146–149, 150, 150, 151, 153, 154, 155, 155, 156, 157, 157, 158, 158, 159, 160, 160, 161, 162, 162, 163, 200, 210, 243, 244, 249, 303
prior
  prior distribution, 24, 107–109, 110, 110, 112, 278–280, 303
  prior estimator, 107
  prior information, 22, 24, 40, 45, 76
privacy, 6, 9, 12, 18, 30–32, 34, 36, 62–66, 170, 171, 174, 175, 179, 182, 185, 187, 192
  differential privacy, 18, 30, 62–66, 171–176, 179, 182, 183, 185, 186, 194, 197, 198, 298
  optimal differential privacy, 192–194, 196, 197
  privacy-preserving learning, 3, 9–10, 17, 46, 61–66, 167, 175, 198
probably approximately correct, 15, 51, 55, 58, 65, 174, 176, 202, 304
probing attack, 11, 31, 40, 45, 48–49, 61, 208, 227, 234, 251, 304
pseudospam attack, 114–116, 119, 121, 122, 126, 128, 128, 129, 131, 243

query, 10, 40, 48, 205, 207, 209, 220, 221, 235–237, 293, 304
  query complexity, 11, 18, 43, 200, 202, 204, 205, 207, 211, 214, 214–217, 220, 228, 229, 231–233, 235–237, 244, 245, 264, 285, 295
  query strategy, 201, 216, 225, 229, 236, 237
  query-based algorithm, 49, 200, 208, 209, 212, 221, 229, 237, 288–290, 293, 302
  query-based evasion problem, 42, 201
  query-based optimization, 201, 202, 211

red herring attack, 51–52
regret, 58, 60, 304
  cumulative regret, 60, 304
  instantaneous regret, 59, 304
  worst-case regret, 60, 304
regret minimization procedure, 60, 304
regularization, 27, 304
reject on negative impact, 55, 114, 120, 128, 129, 131, 243
reject on negative impact defense, 54, 304
residual, 136, 145, 148, 149, 150, 152, 153, 159, 163, 304
  residual subspace, 136, 139, 152
  residual traffic, 159, 160
residual rate, 152, 153, 154, 304
response space, 23, 26, 305
response variable, 23, 305
risk, 26, 171, 176, 177, 180, 184, 185, 188, 305
  empirical risk, 26, 57, 176, 188, 189, 298
  empirical risk minimization, 22, 25–26, 57, 109, 112, 176, 197, 298
robust statistics, 55–58, 305
  breakdown point, 56, 57, 145, 147, 249, 296
  gross-error model, 57, 299
  gross-error sensitivity, 58, 300
  influence function, 56, 57, 249, 300
RONI, vii, 120, 128–130, 131, 132, 243, *see* reject on negative impact

scale invariant, 208, 305
secure learning, 7, 305
security goal, 31, 305
security-sensitive domain, 3, 5, 6, 29, 31, 305
set indicator function, 256, 305
shift invariant, 208, 305
stationarity, 22, 31, 305
support vector machine, 39, 43, 46, 52, 55, 63, 171, 174–179, 182, 184, 185, 187, 192–194, 196–198, 232, 233, 235, 305
  private SVM, 65, 174, 175, 179, 180, 182–188, 191–193

threat model, 13, 31, 306
training, 25–26, 33, 34, 40, 44, 51, 114, 131, 134, 142, 143, 144, 306
  batch learning, 25, 50, 122, 296

https://doi.org/10.1017/9781107338548.016 Published online by Cambridge University Press

training algorithm, 24–26, 32, 106–108, 203, 234, 247, 306
training dataset, 6, 17, 18, 21, 21, 25–27, 30–32, 34, 36, 39, 40, 45–46, 49, 55, 56, 58, 61, 63, 69, 71, 72, 94, 106, 107, 114–116, 119, 120, 128, 134, 162, 171, 197, 234, 247, 278, 281, 306
true-positive rate, 151

unfavorable evaluation distribution, 45, 306
utility, 65, 66, 170, 171, 174–176, 182, 184, 185, 187, 192, 197, 198

VC-dimension, 174–176, 179, 205, 306
vector space, 257, 259
virus detection system, 31, 306

https://doi.org/10.1017/9781107338548.016 Published online by Cambridge University Press