
Index

- A**
- Absolute deviation, 187
 - Account data, 40
 - Account management database, information storage, 40
 - Accuracy ratio (AR), 179
 - AUC, linear relation, 181
 - calculation, example, 181f
 - Acquisition costs, 334t
 - Activation functions, 146
 - Actual fraud, predicted fraud (contrast), 185f
 - Adaptive boosting (Adaboost) procedure, 165–166
 - Adjacency list, 221f, 222
 - Adjacency matrix, 221, 242, 250
 - mathematical representation, 221f
 - Administrative activities (fire incident claims), 12
 - Administrators, experts (collusion), 14
 - Affiliation networks, 267
 - Age
 - default risk, contrast, 60f
 - regression model, 63–64
 - split, entropy (calculation), 139f
 - Agglomerative hierarchical clustering
 - divisive hierarchical clustering, contrast, 94f
 - methods, usage, 95–96
 - Aggregate loss distribution, 329–331
 - description, 328–329
 - indicators, 331f
 - Monte Carlo simulation, 333
 - Alert Type, 292
 - Analysis of variance (ANOVA) test, 143
 - Analytical fraud models
 - backtesting, 302–311
 - calibration, backtesting, 308
 - design/documentation, 311–312
 - life cycle, 280–281
 - performance metric, monitoring, 306t
 - stability, backtesting, 305–307
 - Analytical model life cycle, 280f
 - Analytics, strategic contribution, 337–338
 - Anomaly detection, 343
 - Anticipating effect, 274
 - Anti-fraud steering group, 345
 - Anti-money laundering setting, cash transfers (clustering), 90
 - Approval activities (fire incident claims), 12
 - Approval cycle, absence, 14
 - AR. *See* Accuracy ratio
 - Area under the ROC curve (AUC), 192.
 - See also* Multiclass area under the ROC curve
 - calculation (performance metric), 178
 - Assignment decision. *See* Decision trees
 - Association rule analysis, 88–89
 - Association rules
 - consideration, 89
 - examples, 88
 - Attrition, problem, 338
 - AUC. *See* Area under the ROC curve
 - Autoregressive integrated moving average (ARIMA), 127
 - Average claim value, distribution, 297f
 - Average path length, 212
 - network, 240, 242
- B**
- Backward looking time horizon, 190
 - Backward variable selection, 152f
 - procedure, usage, 151
 - Bagging, 164–165
 - Bankruptcy
 - contrast, 124f
 - filing, 271
 - Base class, 168–169
 - Bayesian methods, 148

- Behavioral characteristics, example, 41–42
 - Behavioral information, 41–42
 - Benford's law, 48–51
 - deviation, 50–51
 - example, 50f
 - Best matching unit (BMU)
 - location, 110
 - weight vector, 109
 - Between-community edges, 259–260
 - quantification, 260
 - weight, sum, 260
 - Betweenness, 227, 239t, 244–246
 - centrality, 246t
 - illustration, 245f
 - recalculation, 262
 - Binary classification, 198–199
 - Binary fraud target (modeling),
 - linear regression (usage), 127
 - Binary link
 - statistics, 253
 - usage, 268–269
 - Binary logistic regression, 169
 - Binary red-flag indicators, 93
 - Binary weight, 216
 - Binomial distribution, usage, 196
 - Binomial test, usage, 309–310
 - Bipartite graph, 267
 - connectivity matrix, 268f
 - example, 267f
 - node types, 268–269
 - usage, 276–277
 - Bipartite networks, 267, 273
 - Bipartite representation, 266–269
 - Birds, clustering (example), 96f
 - dendrogram, 96f
 - BMU. *See* Best matching unit
 - Boosting, 165–166. *See also* Adaptive boosting
 - Bootstrapping, 175, 175f
 - procedure
 - adoption, 307
 - usage, 307
 - Bootstraps, 164–165
 - Bottom-up approaches, 263–264
 - Bottom-up clustering, 263–264
 - Bounding function. *See* Logistic regression
 - Break-point analysis, 343
 - intra-account fraud detection
 - method, 84
 - example, 85f
 - Brier score (BS), measurement, 182
 - Browser-based digital dashboard, usage, 291–292
 - Business policy, 15
 - customer relationship management, example, 14
 - Business rules, set (usage), 290–291
- C**
- C4.5 (decision tree), 137
 - CAIRO matrix, 319
 - Call detail records, example, 20t
 - Cannot-link constraints. *See* Semi-supervised clustering
 - CAP. *See* Cumulative accuracy profile
 - Capital, sufficient level, 326–327
 - CART (decision tree), 137
 - Cascade correlation, 148
 - Case management, 290–295
 - environment, 292
 - Categorical data, 47
 - Categorization, 60–63
 - Chi-squared analysis, usage, 61
 - Centrality. *See* Closeness metrics, 227, 238–246
 - components, 239t
 - CHAID (decision tree), 137
 - Chief analytics officer (CAO), addition, 346
 - Chi-squared analysis, usage, 61
 - Chi-squared distance, calculation, 63
 - Chi-squared distribution, 134
 - test statistic, relationship, 310
 - Chi-squared test statistic, 333
 - Claim
 - amounts, distribution, 297f
 - geographical distribution, 298f
 - enlargement, 299f
 - score, 292
 - value, distribution. *See* Average claim value.
 - ClaimID, 292
 - Classification, 122. *See also* Binary classification
 - accuracy, 176, 189
 - data set, example, 127t
 - error, 176

- measures, dependence, 177
- model, 342
 - calibration, monitoring, 308t
 - SVMs, procedure, 162
 - techniques. *See* Multiclass classification techniques.
- Classifier. *See* Probabilistic relational neighbor
 - cost sensitivity, 199
 - relational neighbor classifier, 233–234
- Click fraud, 6t
- Closed-form solutions, 329
- Closed-loop fraud-management strategy, adoption, 345
- Closeness, 227, 239t, 243–244
 - centrality, 243–244
 - summary, 245t
- Cluster centroids
 - random selection, 105f
 - recalculation, 107f, 108f
 - stability, 104
- ClusterID, 116
- Clustering, 90–116. *See also* Spectral clustering
 - claims, 90
 - constraints, usage, 111–114
 - countries, SOMs (usage), 111f
 - dendrogram, 262f
 - distance metrics, 91–94
 - example. *See* Birds.
 - interpretation, decision trees (usage), 116f
 - screen plot, 97f
 - semi-supervised clustering, must-link/cannot-link constraints, 113f
 - solutions, evaluation/interpretation, 114–116
 - steps, indication, 96f
 - techniques, contrast, 92f, 95f
 - transactions, 90
- Cluster profiling, histograms (usage), 115f
- Clusters
 - analysis, usage. *See* Fraud detection.
 - distances, calculation, 95f
- CNA. *See* Complex network analysis
- Coefficient of determination (R^2), performance metric, 186
- Collective inference algorithms, 227–228, 238, 246–254
- Column-normalized adjacency matrix, 248–249
- Common Neighbor approach, 217
- Communities. *See* Complete communities; Partial communities
 - connection, 244, 246
 - detection. *See* Credit card fraud.
 - fraudulent influence, 265
 - mining, 254–265
 - split (evaluation), modularity Q (usage), 262–263
- Company overview, detection model basis (example), 276t
- Complete communities, 264
 - example, 264f
- Complex network analysis (CNA)
 - benefit, 222
 - usage, 214
- Component plane, usage, 110. *See also* Literacy; Political rights
- Confidence
 - calculation, 89
- Confidence level
 - function, 330
 - indication, 331f
 - selection, decision, 330–331
- CONFIDENTIAL role, 325
- Confusion matrix. *See* Multiclass confusion matrix
 - calculation, 176
 - example, 176t
- Connectance, 226
- Connectivity matrix, 221. *See also* Bipartite graph
 - mathematical representation, 221f
- Constraints. *See* Minimum separation constraint; Must-link constraint
- Constraints, usage, 111–114
- Contextual information, 42
- Contingency table, 67t
- Continuous data, 46–47
- Continuous targets, CAP curve (usage), 187f
- Contracted DT, 283, 285t
- Contractual data, 40
- Convergence, 249–250
- Convolution, 329
- Corporate fraud governance, 343–345

- Corrective measures, complement (preventive actions), 315
 - Corruption, 5t
 - Corruption Perception Index (CPI), 110
 - Cosine measure, basis, 93
 - Cost efficiency, 18
 - Cost-sensitive classifier, 199–200
 - Cost-sensitive cut-off, adoption, 199
 - Cost-sensitive learning, 198–200
 - function, 199–200
 - Counterfeit, 5t
 - Count link statistics, 253
 - Court of Justice (European Union), 326
 - Cramer's V, filter measure (Chi-squared analysis basis), 67
 - Credit card
 - context, time series, 86
 - holder, store (link), 268–269
 - setting, clustering transactions, 90
 - transfer, blocking, 291
 - Credit card fraud, 5t
 - AUC level, 200
 - behavior patterns, 25–26
 - community detection, 259f
 - detection
 - pivot table, example, 82f
 - setting, 284
 - supervised/unsupervised learning, example, 24–26
 - historical cases, 331
 - losses, 9
 - red flags, 58
 - time series data set, 332f
 - toy example, 220f
 - Credit card transaction data
 - example, 219t
 - fields, 24t
 - fraud, network (bipartite graph), 276–277
 - processing, 279
 - Cressey, Donald R., 8
 - Criminal, wrongful category, 3
 - Critical fraud rate, 309
 - Cross-border privacy, regulation, 326
 - Cross-labeled edges, 225
 - Cross-validation. *See* Performance procedure, 174–175
 - Cumulative accuracy profile (CAP), 179
 - curves, usage, 186–187. *See also* Continuous targets.
 - example, 180f
 - Cumulative logistic regression, estimation, 169–170
 - Cumulative notch difference graph, 184f
 - Customer relationship management (CRM) database, information storage, 40
 - Cut-off, 177
 - adoption. *See* Cost-sensitive cut-off.
 - usage, 176f
 - Cyber espionage, 343
- ## D
- Data. *See* Fraud data scientists
 - account data, 40
 - analysis, pie charts. *See* Exploratory data analysis.
 - breach, 343
 - categorical data, 47
 - continuous data, 46–47
 - contractual data, 40
 - correctness, 314
 - data-driven fraud detection, 17–19
 - data item level, outlier detection, 25f
 - data set level, outlier detection, 25f
 - elements, types, 46–47
 - embedding. *See* Unstructured data.
 - external data, access, 325–326
 - features, 271–272
 - importance, 38
 - information, 331–333
 - internal data, access, 319–324
 - missing values, 52–53
 - owners, 316
 - poolers, importance, 41
 - privacy, 317–326
 - protection, 326
 - qualitative, expert-based data, 42–43
 - security issues, 337
 - sources. *See* Publicly available data.
 - merging, 43–44
 - types, 38–43
 - stability, backtesting, 302–304
 - two-step approach, 303
 - standardization, 59
 - stewards, 316–317
 - subscription data, 40
 - surveys, 41

- table, aggregation. *See* Normalized data tables.
- textual data, handling (approaches), 340
- transactional data, 39–40
- unstructured source, 339–340
- Data-driven fraud detection system
 - adoption, 286–287
 - implementation, 2
- Data quality, 314–317
 - indicators
 - design/evaluation, 315–316
 - listings, 316
 - issues, 314–315
 - causes, 315
 - programs/management, 315–317
 - results, action, 316–317
 - root causes, investigation, 316
- Data set. *See* Hierarchical clustering; Linear regression
 - example, 69. *See also* Classification, Impurity; Performance.
 - factual data sets, 272
 - historical data sets, 272
 - observed variance, 70
 - outliers, absence/presence, 84
 - predictive models, development. *See* Skewed data sets.
 - principal component analysis, illustration. *See* Two-dimensional data set.
 - splitting, 172–175
 - transactional data sets, 272–273
 - values, 195f–197f
- Decimal scaling, 59
- Decision boundaries, modeling, 141–142
- Decision tables (DTs), 283–284
 - anomalies, observation, 286
 - contracted DT, 283, 285t
 - expansion, 284t
 - minimization, 285t
 - quadrants, 283
 - tabular representation, usage, 283
 - usage. *See* Rule verification.
- Decision trees, 75, 136–144
 - assignment decision, 137
 - boundaries, 142f
 - building, ClusterID (usage), 116
 - concepts, 136–137
 - error, difference, 168
 - example, 137f
 - forest, creation, 166
 - growth (cessation), validation set (usage), 140f
 - multiclass decision trees, 170
 - power, 144
 - properties, 141–142
 - splitting decision, 137
 - stopping decision, 137
 - usage, 116f
- Decompositional approach, 153f, 163
- Default risk, age (contrast), 60f
- Degree, 228, 239t
 - distribution, 230, 230f
 - illustration, 230f
 - fraudulent degree, 228
 - in-degree, 229
 - legitimate degree, 229
 - out-degree, 229
 - types, 229t
- Degrees of separation, 242
 - theorem. *See* Six degrees of separation theorem.
- Delete, missing value scheme, 52
- Delta-constraints. *See* Semi-supervised clustering
- Dendrogram
 - example, 97f
 - impact, 262
- Denial of service (DoS) attacks, 343
- Density, 228t, 232
 - summary, 233t
- Descriptive analytics, 78
- Descriptive statistics, 51
- Detection model, basis (example), 276t
- Device hacking, 343
- Dicing (OLAP operation), 81
- Dick, Philip K., 23
- Dijkstra’s algorithm, 241
 - illustration, 241f
- Directed network, 229–230
- Direct network features, 271–272
- Dispersion, influence, 258
- Distance metrics, 91–94
- Divergence metric, 182
- Divisive hierarchical clustering
 - agglomerative hierarchical clustering, contrast, 94f
 - initiation, 95–96
- DoS. *See* Denial of service
- Drill-down (OLAP operation), 81

- DTs. *See* Decision tables
- Dyadicity, 226
- Dyadic network, 226–227
- Dynamic characteristics, examples, 41–42
- Dynamic interface/dashboard, usage, 288
- E
- Economic cycle, impact, 340
- Edge. *See* Cross-labeled edges; Hyper-edge; Multi-edge; Self-edge
- betweenness, calculation, 262
- expected probability, 223–225
- ratio. *See* Within-community edges.
- removal, 262
- representation, example, 216
- weight, representation, 217, 222
- EFL. *See* Expected fraud loss
- Ego-centered network (egonet)
- nodes, 218f
- representation, 218
- usage. *See* Social security fraud.
- Eigenvalues, usage, 70
- Eigenvectors, usage, 70
- EL. *See* Expected loss
- Empirical probability distribution, adoption, 328
- Employee numbers, behavioral dynamic characteristic, 41–42
- Encryption techniques, 337
- Ensemble method, 164–168, 188
- bagging, 164–165
- boosting, 165–166
- evaluation, 167–168
- random forests, 166–167
- Entropy, 138–139
- calculation. *See* Age.
- equation, 170
- Gini, contrast, 139f
- nodes, 138
- weighted decrease, 139
- Epochs, 147
- Epsilon-constraint (ϵ -constraint), 113–114. *See also* Semi-supervised clustering
- Error
- estimate, calculation, 177
- type I error, impact, 310
- Error variables
- necessity, 159
- usage, 117–118
- E-sales, rise, 326
- Euclidean distance
- example, 92
- Manhattan distance, contrast, 92f
- metrics, 94–95, 104, 109
- Eulerian path, 210
- European Union (Court of Justice), 326
- Evaluation-related activities (fire incident claims), 12
- Evidence coding, weights, 63–64
- Expected fraud amount, 336
- usage, 290t
- Expected fraud loss (EFL), 122
- calculation, 194
- Expected loss (EL), 327–329, 333
- indication, 331f
- Expected misclassification cost, minimization, 199
- Expert-based approach, 10–11
- Expert-based fraud detection approaches, 2
- Expert-based limits, business knowledge basis, 56
- Exploratory data analysis, pie charts, 49f
- Exploratory statistical analysis, 47–48
- Exposure score, 273
- External data, access, 325–326
- External experts, in-depth assessment (fire incident claims), 12
- External information, 43
- Extracted rules/trees, benchmark, 154
- F
- Factual data sets, 272
- Farness, inverse, 244
- Feature space mapping, example, 160f
- Featurization
- example, 237f
- overview, 254
- process, 254
- unstructured network, mapping, 255t–257t
- FICO score, 41
- Financial loss, avoidance, 14
- Financial shocks, absorption, 316
- Financial statement fraud, AUC level, 200

- Fire incident claims, handling, 12
 - process, example, 13f
- Firewalls, 337
- First digit (frequency distribution description), Benford's law (usage), 50f
- Follower-follower relationships, 215
 - usage. *See* Twitter network.
- Forced claim acceptance, 14
- Forecasting, 338–339
- Forgiving effect, 274
- Forward looking time horizon, 190
- 4-regular graph, 231f
- Fourier transformations, usage, 329
- FP. *See* Fraud percentage
- Fraud
 - alert, 290–295
 - audit trail, 291–292
 - tracking, 291–292
 - amount, 125, 333
 - histogram, 332f
 - usage. *See* Expected fraud amount.
 - analysts, performance metric (usage), 307
 - analytics, economic perspective, 333–336
 - average monthly loss, impact, 333
 - awareness training, 344
 - behavior, volatility, 23
 - call detail records, example, 20t
 - case management, 290–295
 - categories/types, 5t–7t
 - commitment, 258
 - concealment, 3–4
 - confirmation, 22
 - contrast, scatter plot, 185f
 - corrective measures, 15–16
 - cost, 9
 - efficiency, 18
 - crime, organization, 4
 - cycle, 22–23
 - flowchart, 23f
 - data time series, analysis, 328
 - defining, 2–3
 - detection, 10
 - financial impact, 9
 - forecasting, usage, 338–339
 - forms, 4–5
 - investigators, efficiency (evaluation), 301f
 - management, 344
 - mechanism, 10–11
 - comprehension, 11
 - existence, 16–17
 - models. *See* Analytical fraud models.
 - motivational basis, 8–9
 - networks, analysis, 217
 - nonfraud, contrast, 234
 - operational efficiency, 18
 - opportunity, 7f, 8
 - path, existence, 16–17
 - performance benchmarks, 200–201
 - pressure, 7f, 8
 - prevention, 10–12, 22
 - preventive measures, 16
 - probability, 290. *See also* Probability of fraud.
 - calibration, 309
 - propagation, 250
 - questions, 342–344
 - rate. *See* Critical fraud rate.
 - rationalization, 7f, 8
 - revenues, loss, 9
 - scientific perspective, 32–34
 - conclusions, 34
 - scientific publications, statistics, 33f
 - score, assumption, 132
 - setting, association rules, 88–89
 - size, estimation, 9
 - social phenomenon, 9, 222–227
 - treats, examples, 342
 - triangle, 7f
 - uncommonness, 3
 - vulnerability reviews, 344
- Fraud analytics
 - decision trees, usage, 143–144
 - process model, 26–30
 - connections, 26f
- Fraud analytics model
 - characteristics, 28t
 - economical cost, 28t
 - interpretability, 28t
 - operational efficiency, 28t
 - regulatory compliance, 28t
 - representation, 281–286
 - statistical accuracy, 28t
 - traffic light indicator approach, 282–286
- Fraud data scientists, 30–32
 - business comprehension, 32
 - communication/visualization skills, 31–32

- Fraud data scientists (*Continued*)
 creativity, 32
 profile, 33f
 programming expertise, 31
 quantitative skills, 30–31
- Fraud detection, 10–12, 22. *See also*
 Data-driven fraud detection
 big data, 15–17
 capability, 275–276
 cluster analysis, 91f
 domain, benefits, 209
 expert-based approach, 10–11
 improvement, 343
 mission critical application, 314
 model
 basis, example, 276t
 retraining/updating frequency, 23
 OLAP cube, 81f
 PageRank algorithm, impact, 251
 performance benchmarks, 201t
 potential, 296
 power, 23
 precision, 17–18
 process, efficiency (measurement),
 300f
 rule, example, 11
 sample, investigation, 286–290
 scorecard, example, 133f
 system, 94
 data-driven fraud detection
 system, adoption, 286–287
 operation/maintenance,
 335–336
 usage, 287
 techniques, 12–22, 271–272
- Fraud-Fraud label, 226
- Fraud investigation, 22
 activities, 14
 fire incident claims, 12
- Fraud-Legitimate label, 226
- Fraud loss
 capital calculation, 326–333
 Monte Carlo simulation, usage,
 331–333
 evolution/variation, explanation,
 339
 observation, 333–334
 Pareto distributed fraud loss, 333f
- Fraud percentage (FP)
 prediction, 142
 regression tree, usage, 142f
- Fraudsters
 absence, data set values, 195t, 196f
 contact list, 213–214
 exploratory data analysis, 49f
 groups, discovery, 254–265
 oversampling, 191f
 presence, data set values, 197t
 representation, 226
 techniques/tricks, evolution, 4
- Fraudulent bankruptcy, regular
 bankruptcy (contrast), 124f
- Fraudulent degree, summary, 229t
- Fraudulent network, example, 218f
- Fraudulent node, 217
 geodesic paths, 243t
 non-zero values, 250–251
- Fraudulent triangles, 232t
- Frequency
 component, description, 328–329
 domain, 329
- Friend group, 264
 impact, 258
- F*-statistic
 calculation, 143
 intra-cluster similarity, comparison,
 306
- Fully connected network (nodes),
 235
- Fully expanded decision table, 284t
- Funneling effect, 212
- G**
- Gadget fraud, 343
- Garbage in, garbage out (GIGO)
 principle, 315
- Generalized autoregressive conditional
 heteroskedasticity (GARCH), 127
- Genetic algorithms, 148
- Geodesic paths, 227, 239–243. *See also*
 Fraudulent node
 calculations, computation expense,
 239
- Geodesics, number, 238
- Gibbs sampling, 246, 251–252
- GIGO. *See* Garbage in, garbage out
- Gini
 coefficient, 181
 entropy, contrast, 139f
 equation, 170
- Girvan-Newman algorithm, 261–262
 clustering, dendrogram, 262f

- community split, 263
- result, dendrogram (impact), 262
- Global minima, local minima (contrast), 148
- Graph, 215
 - example. *See* (Un)directed graph.
 - extension, bipartite representation, 266–269
 - 4-regular graph, 231f
 - partitioning, 259–260
 - algorithms, 264–265
 - methods, 262–263
 - splitting. *See* Whole graph.
 - theoretic center, 239t, 240
- Graphical outlier detection procedures, 80–83
- Graph theoretic center, 227
- Guilt-by-association, 209, 212, 233
- H**
- Healthcare fraud, 6t
- Heterogeneity, 143
- Heterophilicity, 226
- Heterophobic network, 227
- Hexagonal SOM grid, rectangular SOM grid (contrast), 110f
- Hidden layer, 146
- Hidden neurons
 - number
 - selection, 148
 - variation, 148
 - squares, 150
 - tuning, 148
- Hidden unit activation values, 163–164
 - categorization, clustering (usage), 152
- Hierarchical clustering, 94–97
 - data, scattering plot (usage), 99f
 - data set, 97t
 - initiation. *See* Divisive hierarchical clustering.
 - methods, usage. *See* Agglomerative hierarchical clustering methods.
 - procedures
 - example, 97–104
 - output, 98f–103f
 - techniques, nonhierarchical
 - clustering techniques (contrast), 92f
- Hinton diagram
 - example, 151f
 - inspection, 150
 - usage, 150
- Histograms, usage, 115f
- Historical data sets, 272
- Hit rate, improvement, 287
- Homogeneity, 143
- Homophilic network, 223, 225f
 - real-life example, 224f
- Homophily, 222–227
 - concept, 222
- Hosmer-Lemeshow test, 310
- Hyper-edge, 216
- Hyperplane. *See* Multiple separating hyperplanes
 - defining, 117
 - perpendicular distance, 157–158
- I**
- ICT architecture, 280
- ICT services, 336–337
- Identity theft, 7t
 - frequent contact list, 213f
 - social fraud, 213–214
- If-then classification rules, extraction, 152–153
- If-then rules, 163–164
 - usage, example, 284
- Impurity (calculation), data sets (example), 138f
- In-degree, 229–230
- Independent test set, performance (measurement), 148
- Indirect network features, 272
- Informal economy, average size, 9
- Information. *See* Behavioral information
 - contextual information, 42
 - diffusion, 240
 - exchange, 211–212
 - external information, 43
 - network information, 42
 - storage, 40
 - value filter measure, calculation, 66t
- Information value (IV) filter, weight of evidence basis, 66
- Innocent communities, 264–265
- Insourcing, outsourcing (contrast), 336–338
- Insurance claim
 - filing, 343
 - fraud detection rule, example, 11

- Insurance claim (*Continued*)
 - handling process, internal fraud
 - detection (expert-based approach example), 12–15
 - Insurance fraud, 5t
 - AUC level, 200
 - cost, 9
 - detection, transactions database, 88t
 - example, 125
 - setting, rule set, 285–286
 - Insurance setting, clustering claims, 90
 - Intellectual property (IP) rights, 337
 - Intercept term (β_0), 145
 - Inter-cluster similarity, 306
 - Internal anti-fraud controls, 344
 - Internal data, access, 319–324
 - Internal experts, in-depth assessment (fire incident claims), 12
 - Internet of Things (IoT), 342–344
 - impact, 314
 - Interpretability, 136
 - Intra-account detection method, 85
 - Intra-cluster similarity, *F*-statistic (comparison), 306
 - Intrinsic behavior, description, 237f
 - Intrinsic (local) data feature, 261
 - Intrinsic features, relational features (combination), 275f
 - Intrinsic variables, 236
 - Invalid observations, outlier type, 53
 - IP. *See* Intellectual property
 - Iterative algorithm, usage, 147
 - Iterative bisection, 269
 - example, 261f
 - Iterative classification algorithm, 246, 253
- J**
- Jaccard index, usage, 93
 - Jaccard weight, 217
- K**
- Keep, missing value scheme, 52
 - k*-means clustering, 75, 104–108
 - cluster centroids, recalculation, 107f
 - exercise, output, 116t
 - iteration, 105f
 - cluster centroids, recalculation, 108f
 - iteration observations
 - assignment, 106f
 - reassignment, 108f, 107f
 - original data start, 105f
 - setup, 114
 - k*-nearest neighbors, selection, 193
 - Kolmogorov-Smirnov distance, 181
 - example, 181f
 - Kolmogorov-Smirnov statistic, range, 182
 - Köningsberg bridges, 210, 210f
 - schematic representation, 211f
- L**
- Label-based access control (LBAC), 324–325
 - protection, 325
 - Lagrange multipliers, 118
 - Lagrangian multipliers, usage, 160
 - Lagrangian optimization, usage, 118, 158, 162
 - Laplacian eigenvectors/eigenvalues, 260
 - Lasso regression, 127
 - LBAC. *See* Label-based access control
 - Leading digit, occurrence (mathematical formula), 50
 - Leave node, 143
 - observations, 142
 - Legitimate degree, summary, 229t
 - Legitimate-Legitimate fraud, 226
 - Legitimate node, 217
 - Legitimate triangles, 232t
 - Leniency-related activities (fire incident claims), 12
 - LGF. *See* Loss given fraud
 - Lift curve, 178–179
 - example, 179f
 - Likelihood approach, 194–197
 - Linear decision boundary, 130f
 - Linear kernel, 161
 - Linear programming, 155–156
 - problem, 156
 - Linear regression, 125–127
 - data set, 125t
 - test statistic, 134
 - usage, 127
 - variable selection, 133–136
 - Linear separable case, 156–159
 - Link statistics, types, 253
 - Literacy, component plane (usage), 112

- Local classifier, posterior probabilities, 251–252
- Local minima, 155–156
 global minima, contrast, 148f
- Logistic regression, 129–133. *See also*
 Multiclass logistic regression
 bounding function, 128f
 concepts, 127–129
 formulation, 128
 linear decision boundary, 130f
 model, 128
 estimation, 236, 238
 neural network representation, 145f
 properties, 129–131
 reformulation, 129
 scorecard, building, 131–133
 variable selection, 133–136
- Log-log axes, 230
- Log odds (logit), 129
- Loopy belief propagation, 246, 253
- Lorenz curve, 179
- Loss
 peaks, 327
 random observations, 330
- Loss distribution
 closed-form solution, 329
 derivation, approaches, 329–330
 frequency/severity components, 328–329
 open-form solution, 329–330
- Loss given fraud (LGF), 194
- M**
- MAD. *See* Mean absolute deviation
- Mahalanobis distance
 defining, 83
 performance measure, 182
- Manhattan (city block) distance, 92
 Euclidean distance, contrast, 92f
 metrics, 94–95
- Margin, maximization, 158
- Marital status, good/bad customer
 (contingency table), 67t
- Market basket analysis, 88
- MARS. *See* Multivariate adaptive
 regression splines
- Maximum likelihood, usage, 129
- Mean absolute deviation (MAD), 306,
 307
 definition, 186
- Mean-corrected values, 61
- Mean squared error (MSE), 168, 306,
 307
 calculation, 142
 cost function, optimization, 147
 measure, 186
- MetaCost, introduction, 200
- Milgram, Stanley, 211
- MinCut, 260
- Minima, contrast, 148f
- Minimized decision table, 285t
- Minimum separation (δ) constraint,
 usage, 112–114
- Minkowski distance, 91
- MinMaxCut, 261–265
 metric, 261
- Min/Max standardization, 59
- Minority Report* (Dick), 25
- Misclassifications
 costs, 198t
 error variables, usage, 159
 expected cost, minimization, 199
- Missing values, 52–53
 schemes, 52
 usage, example, 53t
- MLP. *See* Multilayer perceptron
- Modeling, extensions, 338–341
- Mode link statistics, 253
- Modularity Q, 262
 calculation, 263
 maximization, 262f
 usage, 262–263
- Monetary value, 277
- Money laundering, 6t
- Monte Carlo simulation
 impact, 333
 usage, 329–333
- MSE. *See* Mean squared error
- Multiclass area under the ROC curve,
 186
- Multiclass classification techniques,
 168–172
- Multiclass confusion matrix, 183t
- Multiclass decision trees, 170
- Multiclass logistic regression, 168–170
- Multiclass neural networks, 170
- Multiclass problems
 one-versus-all coding, example, 172f
 one-versus-one coding, example,
 171f
- Multiclass support vector machines,
 171–172

- Multiclass targets, performance measures, 182
 - Multicollinearity, 69–70
 - Multidimensional data analysis, 296
 - Multi-edge, 216
 - Multilayer perceptron (MLP) neural network, 145
 - example, 145f
 - hidden layer, 146
 - output layer, 146
 - Multipartite graphs, 269, 269f
 - Multiple separating hyperplanes, 157f
 - Multivariate adaptive regression splines (MARS), 127
 - Multivariate outliers, 54f
 - Multiway splits, 141
 - Must-link constraint, 112–114. *See also* Semi-supervised clustering
- N**
- Negative background information, provision, 114
 - Negative utility, 289
 - Neighborhood, 233
 - behavior, description, 237f
 - impact, metrics, 227–254
 - Neighborhood metrics, 227, 228–238
 - degree, 228–230
 - overview, 228t
 - Networks, 209–222
 - affiliate networks, 267
 - analysis, case study, 270–277
 - average path length, 212, 240, 242
 - bipartite networks, 267
 - central node, 240
 - characteristics/applications, 209–222
 - components, 209, 214–218
 - degree distribution, 230
 - directed network, 229–230
 - dyadicity, 226–227
 - form, 209–222
 - fraudulent network, example, 218f
 - graphical representation, 220
 - heterophobic network, 227
 - homophilic character, 226
 - homophily, 223
 - information, 42
 - mathematical representation, 221f
 - nodes
 - betweenness centrality, 246t
 - closeness/closeness centrality, 245t
 - groups, 258
 - initialization, 277
 - processing, 210
 - relational neighbor probabilities, 234
 - representation, 214f, 219–222. *See also* Ego-centered network (egonet), representation.
 - sample, 229f, 254f, 270f
 - social networks, 211–214
 - Neural networks, 144–155, 188. *See also* Multiclass neural networks
 - black box, opening, 150–155
 - building, 151
 - concepts, 144–146
 - overfitting, prevention, 149
 - reestimation, 151
 - representation, 163f. *See also* Logistic regression.
 - rule extraction
 - decompositional approach, 153f
 - pedagogical approach, 154f
 - shortcomings, 155–156
 - training, 148, 152
 - stopping, validation set (usage), 149f
 - usage, 149
 - Nodes
 - communities, betweenness, 245f
 - connecting paths, number, 242f
 - degree, 228
 - edges, 216
 - fully connected network, 235
 - influence, 266
 - initialization, 277
 - links, 292
 - relationship, 215
 - types, 217
 - integration, 269
 - value, 244
 - Nonfraudsters
 - exploratory data analysis, 49f
 - model, 177
 - undersampling, 191f
 - Nonhierarchical clustering techniques, hierarchical clustering techniques (contrast), 92f
 - Nonlinear regression function, 162
 - Nonlinear SVM classifier, 160–161

- Non-normalized data table, example, 44f
- n -order neighborhood, 227
- Normality, deviations, 57–58
- Normalized data tables, aggregation, 44f
- Normalized weight, 217
- Null hypothesis, 311
- Numerical components, usage, 341
- Numeric weight, 217
- O**
- Observations, 165
 - assignment, 106f
 - classification, necessity, 282–283
 - posterior probability, 199
 - reassignment, 107f, 108f
- Occupational fraud, committing, 8
- OLS. *See* Ordinary least squares
- On-diagonal elements, 182–183
- One-class SVMs, 117–118
 - example, 117f
- One-hop neighborhood, 228
- One-versus-all coding
 - estimates, 171
 - example. *See* Multiclass problems.
 - meta schemes, 172
- One-versus-one coding, 184
 - estimates, 171
 - example, 171f
 - meta schemes, 172
- Online analytical process (OLAP), 26, 296
 - cube. *See* Fraud detection.
 - dicing, 81
 - drill-down, 81
 - functionality, impact, 291–292
 - OLAP-based multidimensional data analysis, 48
 - operations, 79, 81
 - outsourcing, 337–338
 - roll-up, 79, 81
 - slicing, 81
- Online social network sites, usage, 208
- Online transaction processing (OLTP), 39
- Open-form solutions, 329–330
- Operational efficiency, 18, 189
 - variable selection criterion, 136
- Operational risk, 327
 - impact, 328
- Operation costs, 334t
- Opportunity. *See* Fraud
- Optimization
 - problem, 158
 - procedure, 147
- Ordinary least squares (OLS)
 - regression, 126
- Ordinary least squares (OLS), usage, 127
- Organization structure, analytical teams (embedding), 344
- Other People's Money* (Cressey), 8
- Outcome utilities, 335
 - values, 289
- Out-degree, 229–230
- Outlier detection, 25f, 53–57
 - box plots, 55f
 - histogram, 54f
 - procedures
 - graphical outlier detection procedures, 80–83
 - statistical outlier detection procedures, 83–89
 - techniques, 24
 - 2D scatter plot, 80f
 - z-scores, usage, 55t
- Outliers
 - definition, 78
 - multivariate outliers, 54f
 - treatment, 53–57
 - red flags, 57–59
 - t -score, 86–87
 - types, 53
- Out-of-the-loop (O), 319
- Output layer, 146
- Outsourcing
 - analytics, benefits, 336–337
 - insourcing, contrast, 336–338
- Overfitting, 140
 - cause, 149
 - prevention, 149
 - risk, 159
- Overlapping communities, 264–265
 - example, 265f
- Overlapping distributions, SVM classifier (usage), 159f
- Oversampling, 190–192. *See* Fraudsters
 - percentage, level, 193
 - undersampling, combination, 192
 - understanding, 197
- Ownership costs, 334t

P

- PageRank, 246
 - algorithm, 247–251
 - example, 247f, 251t
 - expression, 248
 - extension, 250
 - iterative process, illustration, 249f
 - scores, 250–251
- Parametric Student's *t*-test, usage, 311
- Pareto distributed fraud loss, 333f
- Pareto distribution, 333
- Pareto severity distribution, 333–334
- Partial communities, 264
 - example, 264f
- Partially negative rule, 50
- Pearson correlation
 - basis, 93
 - calculation, 65
 - coefficient, calculation, 185
- Pedagogical approach, 154f
- Peer-group analysis, 85–87
 - advantage, 87–88
 - example, 87f
 - transaction data set, 86t
- Peer pressure, effect, 258
- People-to-people network, 250
 - representation, 266
- Perfectly linearly separable case, SVM classifier (usage), 157f
- Performance
 - benchmarks. *See* Fraud detection.
 - calculation, data set (example), 176t
 - estimation, training/test sample setup (contrast), 173f
 - measurement, 128–189
 - cross-validation, 174f
 - data set, splitting, 172, 174
- Performance metric (PM), 178, 305–306
 - distributions, 307
 - monitoring, 306t
 - statistic test, 307
- PF. *See* Probability of fraud
- Plagiarism, 7t
- PM. *See* Performance metric
- Poisson distribution
 - adoption, 328–329
 - fitting, 331–332
 - monthly fraud frequency, 333f
 - usage, 332
- Poisson frequency distribution, 332–333
- Political rights, component plane, 113
- Polynomial kernel, 161
- Population shift/instability, SSI implication, 304
- Positive utility, 289
- Posterior probabilities, adjustment, 197–198
 - example, 198f
- Post ownership costs, 334t
- Power curve, 179
- Power law, 230
- Predicted fraud, actual fraud (contrast), 185f
- Predictions (calculation), cut-off (usage), 176f
- Predictive analytics, 122
 - classification, 122
 - models, performance measures, 188–189
 - regression, 122
 - types, 122, 338
- Predictive learning, techniques, 338
- Predictive models, evaluation, 172–188
- Pressure. *See* Fraud
- Prudits, 72
 - analysis, 73–74
 - score, 73–74
- Principal components analysis, 68–71
 - illustration, 68f
- Principal components, calculation, 70
- Privacy, 317–326
 - cross-border privacy, regulation, 326
- Probabilistic relational neighbor, 228t classifier, 234–235
 - social network, 235f
- Probability
 - distribution, adoption. *See* Empirical probability distribution.
 - estimates, representation, 252
- Probability of fraud (PF), 194
- Product type variable, coarse
 - classification, 61t, 62t
 - empirical frequencies, options, 62t
 - independence frequencies, options, 63t
 - pivot table, 61t
- Product warranty fraud, 6t
- Propagation algorithm, 273
 - example, 273f
- Pseudo-realistic observations, 332–333
- Publicly available data, sources, 43

- p*-values, 134, 136, 143
 availability, 150
 calculation, student's *t*-distribution (usage), 135f
 computation/representation, 311
- Q**
- Quadratic programming (QP) problem, 118
 solution, Lagrangian optimization (usage), 158
- Qualitative, expert-based data, 42–43
- Quality of Service (QoS) levels, achievement, 276
- R**
- RACI. *See* Responsible Accountable Consulted Informed
- Radial basis function (RBF) kernel, 161
- Random forests, 166–167
 alternative, 167
- Random surfer model, 248
- RASCI. *See* Responsible Approve Support Consult Inform
- RatioCut, 260
 metric, 261
- Rationalization. *See* Fraud
- Ratios, overlap, 69
- RBF. *See* Radial basis function
- REC. *See* Regression error characteristic
- Receiver operating characteristic (ROC) analysis
 introduction, 183–184
 table, 177t
 curve, 177, 173f
 example, 275f
 graph, usage, 181
- Recency frequency and monetary (RFM) variables, 39, 90
 depiction, 82
 framework, usage, 277
 population distribution, 115–116
- Rectangular SOM grid, hexagonal SOM grid (contrast), 109f
- Recursive-partitioning algorithms (RPAs), 136–137
- Red flags, 57–59
 indicators, 93, 94
- Regression, 122
 classifier. *See* Relational logistic regression classifier.
 nonlinear regression function, 162
 SVMs, usage, 161–162, 162f
 time series regression, 339
 trees, 142–143, 164
 usage, 142f
- Regression error characteristic (REC)
 curve, 187
 area, 306, 307
 data, 188t
 example, 188f
- Regression models, 306, 342
 calibration
 backtesting, 310–311
 monitoring, 310t
 performance measures, 185–188
- Regular bankruptcy, fraudulent bankruptcy (contrast), 124f
- Rejected claims, processing, 14
- Relational features, 238t
 intrinsic features, combination, 275f
- Relational logistic regression classifier, 236–233
 social network features, 236f
- Relational neighbor, 228t, 235
 Classifier, 233–234
 social network, example, 233f
 probabilities, 234t
- Relaxation labeling, 246, 253
- Replace (impute), missing value scheme, 52
- Resources, exposure score, 273f
- Responsible Accountable Consulted Informed (RACI) matrix, 318–319
 example, 318f
- Responsible Approve Support Consult Inform (RASCI) matrix, 319
- Return on investment (ROI), 335–336
 formula, calculation, 336
- RFM. *See* Recency frequency and monetary
- Ridge regression, 127
- Ridits, 72–73
 example, 72–73
 scores, 73–74
- Right-tailed test, usage, 309
- Risk
 defining, 327
 operational risk, 327
- RMSE. *See* Root mean squared error
- ROC. *See* Receiver operating characteristic

- ROI. *See* Return on investment
- Roll-up (OLAP operation), 79, 81
- Root mean squared error (RMSE), 186
- RPA. *See* Recursive-partitioning algorithms
- R-squared, 307
- Rule extraction, 150
 - approaches, 163
- Rule set. *See* Insurance fraud
- Rule verification, decision table (usage), 286t

- S**
- Same-labeled edges, 226
- Sample window, variation, 190
- Sampling, 45–46
- SAS Social Network Analysis
 - claim detail investigation, 294f
 - dashboard, 292, 293f
 - link detection, 295f
- SAS Visual Analytics, 296
- Scaling, introduction, 132
- Scatter plot, usage, 99f
- Screen plot, usage (example), 97f
- Section Centrality metrics, 261–262
- Security labels, 324–325
 - components, 324
- Security policies, 324–325
- Segmentation, 74–75
- Seidel's algorithm, 242
- Self-edge, 216
- Self-organizing maps (SOMs), 75, 109–116
 - grids, contrast, 109f
 - usage. *See* Clustering.
 - visualization methods, 110
- Semi-fraudulent triangles, 232t
- Semi-labeled graph, 251–252
- Semi-labeled network, 246
- Semi-supervised clustering
 - delta-constraints (δ -constraints), 114f
 - epsilon-constraints (ϵ -constraints), 114f
 - must-link/cannot-link constraints, 113f
- Shared events, count, 266
- Shortest path, 239
- Sigmoid transformation, 56
- Simple matching coefficient (SMC), usage, 92–94
- Simulated monthly observations, 332–333
- Singular value decomposition (SVD), components, 341
- Sink node, 210–211
- Six degrees of separation theorem, 212
- Skewed data sets
 - oversampling, 190–192
 - predictive models, development, 189–200
 - sample window, variation, 190
 - time window, variation, 190f
 - undersampling, 190–192
- Slicing (OLAP operation), 81
- Small World experiment (Milgram), 211–212
- SMC. *See* Simple matching coefficient
- SMOTE. *See* Synthetic minority oversampling technique
- SMS, usage, 79f
- Social networking sites, 208
- Social networks, 211–214
 - example. *See* Relational neighbor.
 - features. *See* Relational logistic regression classifier.
 - members, connection, 221–222
 - probabilistic relational neighbor classifier, 235f
- Social security fraud, 212–213
 - AUC level, 200
 - egonet, usage, 277f
 - real-life data set, basis, 273f
 - real-life network, degree distribution, 230f
- Sociodemographic information, subscription data source, 40–41
- Soft labeling, 253
- Solvency, behavioral/dynamic characteristic, 41–42
- SOMs. *See* Self-organizing maps
- Source node, 210–211
- Sparse matrix, 221
- Spectral clustering, 260
- Spider constructions, 271. *See also* Tax evasion fraud
- Splitting decision, 137–140. *See also* Decision trees
- Squared error, 187
- Squashing functions, 146
- SSEs. *See* Sum of squared errors
- SSI. *See* System stability index

- Standardization procedures, adoption, 59
- State, term (usage), 283
- Statistical outlier detection procedures, 83–89
- Stepwise logistic regression, usage, 236
- Stopping decision, 140–141. *See also* Decision trees
execution, 170
- Structured data variables, 255t–257t
- Student's *t*-distribution, 136
usage, 135f
- Student's *t*-test, 85, 305
usage. *See* Parametric Student's *t*-test.
- Suboptimal task allocation, 14
- Subscription data, 40
- Sum of squared errors (SSEs), computation, 115
- Supervised model, 343
- Support vector machines (SVMs), 56, 155–164, 188. *See also* Multiclass support vector machines;
One-class SVMs; *v*-SVMs
black box, opening, 163–164
classification SVMs, origins, 156
classifier, 159. *See also* Nonlinear SVM classifier.
building, 161
usage, 159f. *See also* Perfectly linearly separable case.
classifier, neural network representation, 163f
formulation, 162
linear programming, 155–156
problem formulation, 160
usage, 161–162, 162f
- Surveys, 41
- Suspicious activity, call detail records (example), 20t
- Synthetic minority oversampling technique (SMOTE), 192–193
example, 193f
understanding, 197
- System stability index (SSI)
calculation, 303. *See also* Variables.
example, 303t
level, elevation (implication), 304
monitoring, 304t
- T
- Target definition, 123–125
- Tax evasion, 7t
- Tax evasion fraud
example, 122
red flags, 58
spider construction, 124f
- Tax-inspection setting, tax statements (clustering), 90
- TCO. *See* Total cost of ownership
- Telecommunications fraud, 6t
AUC level, 200
- Telecommunications-related fraud, red-flag activities, 58
- Telematics, 343
- Temporal weighing, 273
- Text analytics, 339–341
- Text data, transformation, 341
- Text mapping, 340
- Textual data
ambiguity/complexity, presence, 341
handling, approaches, 343
- Time series regression, 338
- Timestamp, usage, 275
- Time window, variation, 190f
- Total cost of fraud handling, 336
- Total cost of ownership (TCO), 333–335
analysis, goal, 334–335
calculation, costs (example), 334t
- Total degree, summary, 229t
- Trace value, 263
- Traffic light coding procedure, implementation, 306–307
- Traffic light indicator approach, 282–283
representation, 282f
- Training set, 158
- Transactional data, 39–40
sets, 272–273
- Transaction data
example. *See* Credit card.
set. *See* Peer-group analysis.
- Transactions
database, 88t
number, 277
- Transformations, 130
example, 131f
- Transversal standards, development, 345–346
- Traveling Salesman Problem (TSP), 211
- Tree map, display, 296, 302

- Triangles, 228t, 231
types, 232t
- Truncation, z-scores (usage), 57f
- Trusted violators, 8
- t*-score, calculation, 86–87
- TSP. *See* Traveling Salesman Problem
- t*-statistic, value, 85
- t*-test, impact, 311
- Turnover, behavioral/dynamic
characteristic, 41–42
- Twitter network, follower-follower
relationships, 215f
- Two-dimensional data set, principal
component analysis (illustration),
68
- Two-hop path, 243
- Two-stage models, 150
example, 155f
- Type I error, impact, 310
- U
- UL. *See* Unexpected loss
- Uncertainty, description, 327–328
- Uncommon fraud, 3
- Undersampling, 190–192. *See also*
Nonfraudsters
oversampling, combination, 192
understanding, 197
(Un)directed graph, example, 215f
- Unexpected loss (UL), 327–329, 333
absorption, 326–327
calculation, operational risk
(impact), 328
indication, 331f
- Unified distance, (U)-matrix, 111
- Unipartite graphs, 266
example, 266f
- United Kingdom, fraud (impact), 9
- Unstructured data
embedding, 42
sources/types, 339–340
- Unstructured information, 42
- Unstructured network
information, translation, 209
mapping, 255t–257t
- Unsupervised learning, 78
usefulness, 79
- v*-SVMs, 118
- Utility
assignment, 289
components, sum (equivalence),
288–289
negative utility, 289
positive utility, 289
values. *See* Outcome utilities.
- V
- Validation set, usage, 140f, 149f
- Valid observations, outlier type, 53
- Value at Risk (VAR), 127, 330
indication, 331f
- Variables. *See* Intrinsic variables
linear dependency, 65
significance, reference values, 135t
SSI calculation, 305t
subsets, 135f
values, permutation, 167
- Variables selection, 65–68, 150
criteria, 135–136
filters, 65t
performing, 133–134
procedure, 150–151
- Visual analytics, 296–302
- Visual data exploration, 47–48
- W
- WACC. *See* Weighted average cost of
capital
- Ward's distance, 95
- Ward's method, 104
- Web configurable devices, access
(force), 342
- Web of frauds, 223
- Web page
links, 248
rank, 247
- Weight. *See* Binary weight; Jaccard
weight; Normalized weight;
Numeric weight
list, 221f, 222
matrix
edge weight expression, 222
mathematical representation, 221f
- Weighted average cost of capital
(WACC), 334–335
- Weight learning, 147–149
- Weight of evidence (WOE)
calculations, 64t
coding, 63–64
usage, 131–132

- Weight parameter vector (β),
 - closed-form formula, 126
 - Weight regularization, 149
 - White-box analytical technique, 153
 - White-box model, 144
 - Whole graph, splitting, 259
 - WITH CHECK option, 324
 - Within-community edges, 263
 - between-community edges, ratio, 259–260
 - World Wide Web technology stack, 342
 - Wrongful, human category, 3
- Z**
- Z-scores
 - calculation, 55, 84
 - standardization, 59
 - usage. *See* Outlier detection; Truncation.

<http://www.pbookshop.com>

<http://www.pbookshop.com>