# Artificial Intelligence Model Extension

## Table of Contents


1. Supervised Learning – Regression
2. Supervised Learning – Classification
3. Unsupervised – Clustering
4. Unsupervised – Dimensionality Reduction
5. Unsupervised – Anomaly Detection
6. Deep Learning – Computer Vision
7. Deep Learning – NLP
8. Deep Learning – Time Series
9. Reinforcement Learning
10. Ensemble Methods
11. Specialized Models
12. AutoML & Optimization
13. Probabilistic Models
14. Causal Inference
15. Survival Analysis
16. Transfer Learning
17. Edge AI & Model Compression
18. Federated & Distributed Learning
19. File Extension Reference Guide
20. Best Practices for Model Storage


## 1 — SUPERVISED LEARNING (REGRESSION)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| **Linear Regression** | Predict continuous values | OLS, Ridge, Lasso, ElasticNet | .pkl, .joblib, .h5, .json, .pmml | Sales forecasting, price prediction | Beginner |
| **Polynomial Regression** | Non-linear relationships | Polynomial features + Linear models | .pkl, .joblib, .sav | Growth curves, physical phenomena | Beginner |
| **Decision Tree Regression** | Non-linear, interpretable | CART, scikit-learn DecisionTreeRegressor | .pkl, .joblib, .dot, .xml | Risk assessment, medical diagnosis | Intermediate |
| **Random Forest Regression** | Robust ensemble predictions | scikit-learn RandomForestRegressor | .pkl, .joblib, .h5 | Real estate valuation, stock prices | Intermediate |
| **Gradient Boosting Regression** | High-accuracy predictions | XGBoost, LightGBM, CatBoost | .pkl, .bin, .txt, .cbm, .json | Kaggle competitions, demand forecasting | Intermediate |
| **Support Vector Regression (SVR)** | Complex non-linear patterns | scikit-learn SVR | .pkl, .joblib, .sav | Financial modeling, energy consumption | Intermediate |
| **Neural Network Regression** | Deep learning for complex data | MLP, Deep Neural Networks | .h5, .pb, .pth, .onnx, .pt | Complex pattern recognition | Advanced |
| **Bayesian Ridge Regression** | Probabilistic regression | scikit-learn BayesianRidge | .pkl, .joblib | Uncertainty quantification | Intermediate |
| **Huber Regression** | Robust to outliers | scikit-learn HuberRegressor | .pkl, .joblib | Noisy data, robust modeling | Intermediate |
| **Quantile Regression** | Predict conditional quantiles | statsmodels, scikit-learn | .pkl, .joblib | Risk management, interval predictions | Advanced |

| | | | | | |
|---|---|---|---|---|---|
| **Generalized Additive Models (GAM)** | Flexible non-linear modeling | pyGAM, statsmodels | .pkl, .joblib | Ecological modeling, epidemiology | Advanced |

---

## 2 — SUPERVISED LEARNING (CLASSIFICATION)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Logistic Regression** | Binary/multi-class classification | scikit-learn LogisticRegression | .pkl, .joblib, .pmml, .json | Credit scoring, medical diagnosis | Beginner |
| **K-Nearest Neighbors (KNN)** | Instance-based classification | scikit-learn KNeighborsClassifier | .pkl, .joblib, .sav | Recommendation systems, pattern recognition | Beginner |
| **Naive Bayes** | Probabilistic classification | GaussianNB, MultinomialNB, BernoulliNB | .pkl, .joblib, .json | Spam detection, sentiment analysis | Beginner |
| **Decision Tree Classifier** | Rule-based classification | scikit-learn DecisionTreeClassifier | .pkl, .joblib, .dot, .json, .xml | Customer segmentation, fraud detection | Intermediate |
| **Random Forest Classifier** | Ensemble classification | scikit-learn RandomForestClassifier | .pkl, .joblib, .h5, .json | Healthcare, finance, e-commerce | Intermediate |

| | | | | | |
|---|---|---|---|---|---|
| **Gradient Boosting Classifier** | High-performance classification | XGBoost, LightGBM, CatBoost | .pkl, .model, .bin, .txt, .cbm, .json | CTR prediction, fraud detection | Intermediate |
| **AdaBoost** | Adaptive boosting | scikit-learn AdaBoostClassifier | .pkl, .joblib | Face detection, text classification | Intermediate |
| **Extra Trees** | Extremely randomized trees | scikit-learn ExtraTreesClassifier | .pkl, .joblib | Feature importance, fast training | Intermediate |
| **Support Vector Machine (SVM)** | Maximum margin classification | scikit-learn SVC | .pkl, .joblib, .sav, .pmml | Image classification, bioinformatics | Intermediate |
| **Linear Discriminant Analysis (LDA)** | Linear decision boundary | scikit-learn LinearDiscriminantAnalysis | .pkl, .joblib | Face recognition, pattern classification | Intermediate |
| **Multi-layer Perceptron (MLP)** | Neural network classifier | scikit-learn MLPClassifier, TensorFlow | .pkl, .h5, .pb, .pth, .onnx | Deep classification, complex patterns | Advanced |

---

### 3 — UNSUPERVISED LEARNING (CLUSTERING)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| **K-Means** | Partition into K clusters | scikit-learn KMeans, MiniBatchKMeans | .pkl, .joblib, .json | Customer segmentation, image compression | Beginner |
| **Hierarchical Clustering** | Tree-based clustering | AgglomerativeClustering | .pkl, .joblib, .newick | Taxonomy, gene analysis | Intermediate |
| **DBSCAN** | Density-based clustering with noise handling | scikit-learn DBSCAN | .pkl, .joblib | Spatial clustering, anomaly detection | Intermediate |
| **HDBSCAN** | Hierarchical DBSCAN for variable density | hdbscan library | .pkl, .joblib | Complex density-based clustering | Advanced |
| **Gaussian Mixture Models (GMM)** | Probabilistic soft clustering | scikit-learn GaussianMixture | .pkl, .joblib, .json | Image segmentation, speech recognition | Intermediate |
| **Mean Shift** | Centroid-based clustering without fixed K | scikit-learn MeanShift | .pkl, .joblib | Image segmentation, object tracking | Intermediate |
| **Spectral Clustering** | Graph-based clustering | scikit-learn SpectralClustering | .pkl, .joblib | Community detection, image segmentation | Advanced |

---

## 4 — UNSUPERVISED LEARNING (DIMENSIONALITY REDUCTION)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Principal Component Analysis (PCA)** | Linear feature reduction | scikit-learn PCA | .pkl, .joblib, .npz | Data compression, noise reduction | Beginner |

| | | | | | |
|---|---|---|---|---|---|
| **t-SNE** | Non-linear visualization | scikit-learn TSNE | .pkl, .joblib, .npy | High-dimensional visualization | Intermediate |
| **UMAP** | Fast non-linear visualization | umap-learn | .pkl, .joblib | Large-scale visualization | Intermediate |
| **Autoencoders** | Neural feature compression | TensorFlow, PyTorch | .h5, .pb, .pth, .pt, .onnx | Feature learning, compression | Advanced |
| **Independent Component Analysis (ICA)** | Signal separation | scikit-learn FastICA | .pkl, .joblib | Signal processing, blind source separation | Advanced |

---

## 5 — UNSUPERVISED LEARNING (ANOMALY DETECTION)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Isolation Forest** | Tree-based anomaly detection | scikit-learn IsolationForest | .pkl, .joblib | Fraud detection, network security | Intermediate |
| **One-Class SVM** | Novelty detection in high dimensions | scikit-learn OneClassSVM | .pkl, .joblib, .sav | Outlier detection, anomaly scoring | Intermediate |
| **Local Outlier Factor (LOF)** | Density-based anomaly detection | scikit-learn LocalOutlierFactor | .pkl, .joblib | Local outlier detection | Intermediate |
| **Elliptic Envelope** | Gaussian distribution–based outlier detection | scikit-learn EllipticEnvelope | .pkl, .joblib | Data cleaning, anomaly isolation | Intermediate |
| **Autoencoders for Anomaly Detection** | Reconstruction error detection | TensorFlow, PyTorch | .h5, .pb, .pth, .pt | Complex anomaly patterns | Advanced |

---

## 6 — DEEP LEARNING (COMPUTER VISION)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Convolutional Neural Networks (CNN)** | Image classification | TensorFlow, PyTorch, Keras | .h5, .pb, .pth, .pt, .onnx, .tflite | Medical imaging, autonomous vehicles | Advanced |
| **ResNet (Residual Networks)** | Deep networks with skip connections | ResNet-50/101/152 | .h5, .pb, .pth, .pt, .onnx | ImageNet, transfer learning | Advanced |
| **Inception (GoogLeNet)** | Multi-scale feature extraction | InceptionV3, InceptionResNetV2 | .h5, .pb, .pth, .onnx | Efficient image classification | Advanced |
| **MobileNet** | Lightweight mobile models | MobileNetV2, MobileNetV3 | .h5, .tflite, .pb, .onnx | Edge/mobile deployment | Advanced |
| **EfficientNet** | Compound scaling for efficiency | EfficientNet B0–B7 | .h5, .pb, .pth, .onnx | High-accuracy, efficient CV models | Advanced |
| **Vision Transformers (ViT)** | Transformer applied to images | Hugging Face transformers | .bin, .safetensors, .pth, .onnx | Advanced image classification | Expert |
| **YOLO (You Only Look Once)** | Real-time object detection | YOLOv5, YOLOv8, Ultralytics | .pt, .onnx, .tflite, .weights, .cfg | Surveillance, autonomous driving | Advanced |
| **R-CNN / Faster R-CNN** | Region-based detection | Detectron2 | .pth, .pkl, .yaml | High-precision object detection | Advanced |
| **U-Net** | Medical image segmentation | TensorFlow, PyTorch | .h5, .pb, .pth, .pt | Biomedical segmentation | Advanced |

| | | | | | |
|---|---|---|---|---|---|
| **Mask R-CNN** | Instance segmentation | Detectron2 | .pth, .pkl, .yaml | Object instance-level segmentation | Advanced |
| **Semantic Segmentation Models** | Pixel-wise classification | DeepLab, FCN | .h5, .pb, .pth, .pt | Autonomous driving, medical | Advanced |
| **Siamese Networks** | Similarity learning | PyTorch, TensorFlow | .h5, .pb, .pth, .pt | Face verification, signature verification | Advanced |
| **Facial Recognition Models** | Face identification/verification | FaceNet, ArcFace, DeepFace | .h5, .pb, .pth, .pt, .onnx | Authentication, biometrics | Advanced |
| **Pose Estimation** | Human pose detection | OpenPose, MediaPipe | .pb, .pth, .tflite, .onnx | Sports, AR/VR | Advanced |
| **GANs (Generative Adversarial Networks)** | Image generation | StyleGAN, DCGAN, CycleGAN | .h5, .pb, .pth, .pt, .pkl | Art, synthetic data, augmentation | Expert |
| **Diffusion Models** | High-quality generative models | Stable Diffusion, DALL-E | .pth, .safetensors, .ckpt, .bin | Text-to-image generation | Expert |
| **Variational Autoencoders (VAE)** | Probabilistic generation | TensorFlow, PyTorch | .h5, .pb, .pth, .pt | Image generation, anomaly detection | Advanced |
| **Super Resolution (SRGAN/ESRGAN)** | Image upscaling | ESRGAN, Real-ESRGAN | .pth, .pb, .h5 | Image enhancement | Advanced |
| **Optical Character Recognition (OCR)** | Text extraction from images | Tesseract, EasyOCR, TrOCR | .traineddata, .pth, .bin | Document processing, automation | Advanced |

## 7 — DEEP LEARNING (NATURAL LANGUAGE PROCESSING)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Word Embeddings** | Dense word representations | Word2Vec, GloVe, FastText | .bin, .vec, .txt, .model | Text similarity, search engines | Intermediate |
| **Bag of Words (BoW)** | Simple text vectorization | CountVectorizer, TfidfVectorizer | .pkl, .joblib, .npz | Basic text classification | Beginner |
| **Long Short-Term Memory (LSTM)** | Sequence modeling, long-term dependencies | TensorFlow, PyTorch | .h5, .pb, .pth, .pt, .onnx | Language modeling, translation | Advanced |
| **Gated Recurrent Unit (GRU)** | Reduced-complexity sequence modeling | TensorFlow, PyTorch | .h5, .pb, .pth, .pt, .onnx | Chatbots, predictive typing | Advanced |
| **BERT (Bidirectional Encoder)** | Contextual embeddings | bert-base, bert-large, DistilBERT | .bin, .safetensors, .pth, .h5, .onnx | Text classification, QA, NER | Advanced |
| **GPT (Generative Pre-trained Transformer)** | Autoregressive text generation | GPT-2, GPT-3, GPT-4 | .bin, .safetensors, .pth, .ckpt | Chatbots, content generation | Expert |
| **T5 (Text-to-Text Transfer Transformer)** | Unified text tasks (translate, QA, summarize) | Hugging Face T5 | .bin, .safetensors, .pth, .h5 | Summarization, translation | Advanced |
| **RoBERTa** | Optimized BERT variant | Hugging Face RoBERTa | .bin, .safetensors, .pth, .onnx | Improved NLP performance | Advanced |
| **Named Entity Recognition (NER)** | Entity extraction | spaCy, BERT-NER, Flair | .pkl, .bin, .pth, .model | Information extraction | Advanced |

| | | | | | |
|---|---|---|---|---|---|
| **Sentiment Analysis Models** | Opinion mining | VADER, TextBlob, BERT-based | .pkl, .bin, .pth, .h5 | Social analytics, reviews | Intermediate |
| **Question Answering Models** | Extractive QA | BERT-QA, RoBERTa-QA | .bin, .safetensors, .pth, .h5 | Chatbots, customer support | Advanced |
| **Text Summarization Models** | Generate summaries | BART, T5, Pegasus | .bin, .safetensors, .pth, .h5 | Document automation | Advanced |
| **Machine Translation Models** | Language translation | MarianMT, mBART, M2M-100 | .bin, .safetensors, .pth, .spm | Globalization, localization | Advanced |
| **Large Language Models (LLM)** | Large-scale reasoning & generation | GPT-4, Claude, LLaMA, Gemini | .bin, .safetensors, .pth, .gguf, .ckpt | Chatbots, assistants, agents | Expert |
| **Speech Recognition Models** | Audio → text | Whisper, Wav2Vec2, DeepSpeech | .pb, .pth, .bin, .onnx, .tflite | Transcription, voice assistants | Advanced |
| **Text-to-Speech (TTS)** | Speech synthesis | Tacotron, WaveNet, FastSpeech | .pth, .pb, .h5, .onnx | IVR, audiobooks, assistants | Advanced |
| **Topic Modeling** | Discover hidden topics | LDA, NMF, BERTopic | .pkl, .joblib, .model | Document clustering | Intermediate |

---

## 8 — DEEP LEARNING (TIME SERIES)

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **ARIMA / SARIMA** | Statistical forecasting | statsmodels ARIMA | .pkl, .joblib, .json | Traditional forecasting, finance | Intermediate |
| **Prophet** | Business time-series forecasting | fbprophet | .pkl, .json | Demand forecasting, KPI prediction | Intermediate |

| | | | | | |
|---|---|---|---|---|---|
| **LSTM for Time Series** | Sequential deep forecasting | TensorFlow, PyTorch | .h5, .pb, .pth, .pt, .onnx | Stock prediction, IoT signals | Advanced |
| **Temporal Convolutional Networks (TCN)** | CNN-based sequence modeling | PyTorch | .pth, .pt, .onnx | High-speed time-series tasks | Advanced |
| **Transformer for Time Series** | Attention-based forecasting | Temporal Fusion Transformer | .pth, .pt, .bin, .ckpt | Long-range dependency forecasting | Advanced |
| **N-BEATS** | Deep neural forecasting model | Darts, PyTorch | .pth, .pt, .pkl | Pure deep learning forecasting | Advanced |
| **DeepAR** | Probabilistic time-series forecasting | GluonTS | .params, .pkl, .json | Retail forecasting, uncertainty modeling | Advanced |

---

## 9 — REINFORCEMENT LEARNING

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Q-Learning** | Value-based learning for discrete states | Custom implementations | .pkl, .npy, .json | Gridworld, simple control systems | Advanced |
| **Deep Q-Networks (DQN)** | Deep value-based reinforcement learning | TensorFlow, PyTorch | .h5, .pb, .pth, .pt | Atari games, robotics simulation | Expert |
| **Proximal Policy Optimization (PPO)** | Stable policy-gradient optimization | Stable-Baselines3, OpenAI | .zip, .pkl, .pth, .pb | Industry-standard RL, robotics | Expert |
| **Soft Actor-Critic (SAC)** | Entropy-regularized policy optimization | Stable-Baselines3 | .zip, .pkl, .pth | Continuous control tasks | Expert |

| | | | | | |
|---|---|---|---|---|---|
| **Actor-Critic (A2C / A3C)** | Combined value + policy learning | Stable-Baselines3, RLlib | .zip, .pkl, .pth, .pb | Games, multi-agent systems | Expert |

---

## 10 — ENSEMBLE METHODS

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Bagging** | Reduce variance using bootstrap aggregation | Random Forest, BaggingClassifier | .pkl, .joblib | Stable predictions in noisy datasets | Intermediate |
| **XGBoost** | Extreme gradient boosting | xgboost library | .model, .json, .bin, .pkl | Kaggle competitions, structured data | Intermediate |
| **LightGBM** | Fast gradient boosting | lightgbm library | .txt, .bin, .pkl, .model | Large datasets, real-time predictions | Intermediate |
| **CatBoost** | Handling categorical features natively | catboost library | .cbm, .bin, .pkl, .json | CTR prediction, tabular ML | Intermediate |
| **Stacking** | Multi-model ensemble | mlxtend, scikit-learn | .pkl, .joblib | Combine diverse models for higher accuracy | Advanced |
| **Voting** | Aggregate predictions (hard/soft voting) | VotingClassifier | .pkl, .joblib | Simple and effective ensemble method | Intermediate |

---

## 11 — SPECIALIZED AI/ML MODELS

| Model Type | Primary Use Case | Key Algorithms | Common File | Industry Applications | Skill Level |
|---|---|---|---|---|---|

| | | / Frameworks | Extensions | | |
|---|---|---|---|---|---|
| **Recommendation Systems (Collaborative Filtering)** | User–item pattern prediction | Surprise, implicit, ALS | .pkl, .model, .bin | E-commerce, OTT platforms | Intermediate |
| **Matrix Factorization** | Latent factor extraction | SVD, NMF, SVD++ | .pkl, .npy, .model | Netflix Prize, recommender engines | Advanced |
| **Neural Collaborative Filtering (NCF)** | Deep learning–based recommendations | TensorFlow, PyTorch | .h5, .pb, .pth, .pt | Modern recommender systems | Advanced |
| **Deep Factorization Machines (DeepFM)** | Automatic feature interaction learning | DeepFM, xDeepFM | .h5, .pb, .pth | Ad-click prediction, personalization | Advanced |
| **Graph Convolutional Networks (GCN)** | Node classification on graphs | PyTorch Geometric, DGL | .pth, .pt, .pkl, .bin | Social networks, chemistry | Expert |
| **Graph Attention Networks (GAT)** | Attention over graph nodes | PyTorch Geometric | .pth, .pt, .pkl | Drug discovery, knowledge graphs | Expert |
| **GraphSAGE** | Inductive graph learning | PyTorch Geometric | .pth, .pt, .pkl | Large-scale graph ML | Expert |
| **Knowledge Graph Embeddings** | Learn entity/relation representations | TransE, DistMult, ComplEx | .pkl, .npy, .pth | Knowledge bases, semantic search | Expert |
| **CLIP (Multimodal)** | Image–text alignment | OpenAI CLIP | .pth, .pt, .bin, .safetensors | Search, tagging, vision-language | Expert |
| **ALIGN** | Large-scale multimodal learning | Google Research | .pb, .pth, .ckpt | Visual understanding | Expert |

| | | | | | |
|---|---|---|---|---|---|
| **Flamingo** | Few-shot multimodal reasoning | DeepMind | .pth, .safetensors, .ckpt | VQA, multimodal assistants | Expert |
| **GPT-4 Vision** | Multimodal text+image understanding | OpenAI API | API-based internal files | Image reasoning, agents | Expert |

---

## 12 — AUTOML & OPTIMIZATION

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Grid Search** | Exhaustive hyperparameter search | scikit-learn GridSearchCV | .pkl, .joblib, .json | Small hyperparameter spaces | Beginner |
| **Bayesian Optimization** | Probabilistic hyperparameter tuning | Optuna, Hyperopt | .pkl, .db, .json | Efficient tuning for ML/DL models | Advanced |
| **Neural Architecture Search (NAS)** | Automated model architecture creation | AutoKeras, NAS-Bench | .h5, .pb, .json | Discovering optimal deep-learning architectures | Expert |
| **AutoML Frameworks** | End-to-end ML automation | Auto-sklearn, AutoGluon, H2O | .pkl, .zip, .bin | Rapid model building for enterprises | Intermediate |
| **SHAP** | Explainable AI via feature attribution | shap library | .pkl, .json | Interpretability for ML models | Intermediate |
| **LIME** | Local interpretability for predictions | lime library | .pkl, .json | Instance-level explanations | Intermediate |
| **Integrated Gradients** | Deep learning interpretability | Captum (PyTorch) | .pth, .pt | Explainable neural networks | Advanced |

---

## 13 — PROBABILISTIC MODELS

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Bayesian Networks** | Probabilistic graphical modeling | pgmpy, PyMC | .pkl, .bif, .xml, .json | Causal reasoning, diagnostics | Advanced |
| **Hidden Markov Models (HMM)** | Sequential state modeling | hmmlearn | .pkl, .json | Speech recognition, bioinformatics | Advanced |
| **Gaussian Processes (GP)** | Non-parametric regression | GPy, scikit-learn | .pkl, .json, .npy | Uncertainty-aware forecasting | Advanced |
| **Markov Chain Monte Carlo (MCMC)** | Bayesian posterior sampling | PyMC3, Stan | .nc, .pkl, .stan | Scientific modeling, Bayesian inference | Expert |
| **Probabilistic Programming** | Flexible statistical modeling | PyMC3, TensorFlow Probability | .nc, .pkl, .pb | Custom probabilistic ML models | Expert |

---

## 14 — CAUSAL INFERENCE

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Propensity Score Matching (PSM)** | Treatment effect estimation | DoWhy, CausalML | .pkl, .json | A/B testing, healthcare, policy impact | Advanced |
| **Causal Forests** | Estimate heterogeneous treatment effects | EconML, CausalML | .pkl, .joblib | Personalized medicine, marketing | Expert |
| **Structural Causal Models (SCM)** | Causal graph modeling | DoWhy, pgmpy | .pkl, .dot, .gml | Root-cause analysis, causal reasoning | Expert |

---

## 15 — SURVIVAL ANALYSIS

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Cox Proportional Hazards Model** | Time-to-event modeling | lifelines | .pkl, .json | Medical prognosis, churn prediction | Advanced |
| **Kaplan–Meier Estimator** | Non-parametric survival curves | lifelines | .pkl, .json, .csv | Clinical trials, retention analysis | Intermediate |
| **Deep Survival Models** | Neural survival analysis | pycox, DeepSurv | .pth, .pt, .h5 | Healthcare survival, risk prediction | Expert |

---

## 16 — TRANSFER LEARNING

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Pre-trained Models** | Use existing models for downstream tasks | Hugging Face Hub, TensorFlow Hub | .bin, .safetensors, .pb, .pth | Quick deployment, reduced training cost | Intermediate |
| **Fine-Tuning** | Adapt pre-trained models to new domains | Hugging Face Trainer | .bin, .safetensors, .pth, .h5 | Domain-specific NLP/CV tasks | Advanced |
| **Domain Adaptation** | Transfer across different data distributions | DANN, CORAL | .pth, .pt, .pkl | Cross-domain applications | Expert |
| **Few-Shot Learning** | Learn from very small datasets | Prototypical Networks, MAML | .pth, .pt | Low-data use cases | Expert |

---

## 17 — EDGE AI & MODEL COMPRESSION

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Quantization** | Reduce model precision for faster inference | TensorFlow Lite, PyTorch | .tflite, .onnx, .pth | Mobile/edge deployment | Advanced |
| **Pruning** | Remove unnecessary weights | TensorFlow Model Optimization Toolkit | .h5, .pb, .pth | Reduce model size & latency | Advanced |
| **Knowledge Distillation** | Compress model using teacher–student learning | Custom implementations | .h5, .pb, .pth, .pt | Lightweight high-performance models | Advanced |
| **ONNX Conversion** | Framework-agnostic model deployment | ONNX Runtime | .onnx | Cross-platform inference | Intermediate |

---

## 18 — FEDERATED & DISTRIBUTED LEARNING

| Model Type | Primary Use Case | Key Algorithms / Frameworks | Common File Extensions | Industry Applications | Skill Level |
|---|---|---|---|---|---|
| **Federated Learning** | Decentralized privacy-preserving training | TensorFlow Federated, PySyft | .h5, .pb, .pth, .pt | Healthcare, finance, on-device learning | Expert |
| **Distributed Training** | Multi-GPU / multi-node deep learning | Horovod, PyTorch DDP | .pth, .pt, .pb, .ckpt | Large-scale deep learning workloads | Advanced |

---

## 19 — FILE EXTENSION REFERENCE GUIDE

### A. Python / Scikit-learn Model Formats

| Extension | Description / Use |
|---|---|

| | |
|---|---|
| **.pkl** | Pickle format (Python object serialization) |
| **.joblib** | Efficient serialization for NumPy-heavy models |
| **.sav** | Alternative model save format |
| **.npz** | Compressed NumPy arrays |
| **.npy** | NumPy array format |
| **.json** | Model metadata/config in JSON |

---

**B. TensorFlow / Keras Formats**

| Extension | Description / Use |
|---|---|
| **.h5** | HDF5 (weights + architecture) |
| **.pb** | Protocol Buffer SavedModel |
| **.tflite** | TensorFlow Lite model (mobile/edge) |
| **.saved_model** | Directory-based SavedModel |
| **.ckpt** | TensorFlow checkpoint files |

---

**C. PyTorch Formats**

| Extension | Description / Use |
|---|---|
| **.pth** | PyTorch model weights |
| **.pt** | PyTorch model/tensor |
| **.tar** | Compressed checkpoint |
| **.ckpt** | Training checkpoint |

---

**D. Hugging Face Formats**

| Extension | Description / Use |
|---|---|
| **.bin** | Binary model weights |
| **.safetensors** | Secure/optimized serialization |

| | |
|---|---|
| **.json** | Config files |

---

**E. Tree-Based Model Formats**

| Extension | Description / Use |
|---|---|
| **.model** | XGBoost model |
| **.json** | XGBoost JSON dump |
| **.txt** | LightGBM text model |
| **.bin** | LightGBM binary model |
| **.cbm** | CatBoost binary model |

---

**F. NLP-Specific Formats**

| Extension | Description / Use |
|---|---|
| **.vec** | Word vectors (Word2Vec, GloVe) |
| **.bin** | FastText binary model |
| **.model** | Gensim model |
| **.spm** | SentencePiece tokenizer model |
| **.traineddata** | Tesseract OCR file |

---

**G. YOLO Formats**

| Extension | Description / Use |
|---|---|
| **.weights** | Darknet YOLO weights |
| **.cfg** | YOLO configuration |
| **.pt** | YOLO PyTorch model |
| **.onnx** | ONNX export |

---

**H. Interoperability Formats**

| Extension | Description / Use |
|---|---|
| **.onnx** | Open Neural Network Exchange (cross-framework) |

| | |
|---|---|
| **.pmml** | Predictive Model Markup Language |
| **.coreml** | Apple Core ML model |

---

**I. Configuration Files**

| Extension | Description / Use |
|---|---|
| **.yaml / .yml** | Configuration files |
| **.json** | Metadata/config |
| **.xml** | XML format |
| **.ini** | Initialization/config format |

---

**J. Graph & Network Files**

| Extension | Description / Use |
|---|---|
| **.dot** | Graphviz |
| **.gml** | Graph Modeling Language |
| **.graphml** | GraphML structure |
| **.newick** | Newick tree format |

---

**K. Bayesian / Statistical Formats**

| Extension | Description / Use |
|---|---|
| **.nc** | NetCDF (PyMC3 traces) |
| **.stan** | Stan models |
| **.bif** | Bayesian network structure |
| **.pkl** | Stored samples/models |

---

**L. Database / Storage Formats**

| Extension | Description / Use |
|---|---|

| | |
|---|---|
| **.db** | SQLite database (Optuna) |
| **.zip** | Compressed model bundles |

---

**M. Large Language Model Formats**

| Extension | Description / Use |
|---|---|
| **.gguf** | GPT Unified Format (llama.cpp) |
| **.ggml** | Legacy GGML format |

---

## 20 — BEST PRACTICES FOR MODEL STORAGE

**A. Recommended Directory Structure (Version Control)**

| Folder / File | Purpose |
|---|---|
| models/model_v1.0.pkl | Version 1 of model |
| models/model_v1.1.pkl | Version 1.1 of model |
| models/config.yaml | Configurations for training/inference |
| models/metadata.json | Metadata: metrics, parameters, timestamps |

---

**B. Recommended Formats by Use Case**

| Use Case | Recommended Format | Reason |
|---|---|---|
| **Production (Python)** | .joblib, .pkl | Fast serialization |
| **Deep Learning (TensorFlow)** | .pb, .h5 | Standard TF formats |
| **Deep Learning (PyTorch)** | .pth, .pt | Native PyTorch formats |
| **Cross-platform deployment** | .onnx | Framework interoperability |
| **Mobile / Edge deployment** | .tflite, .onnx | Optimization for size & speed |
| **Model Registry** | .pkl + .json | Model + metadata management |
| **Large Models** | .safetensors, .bin | Efficient & safe storage |

---

**C. File Size Recommendations**

| Model Size | Recommended Format | Notes |
|---|---|---|
| **Small models (<100MB)** | .pkl, .joblib | Fast load times |
| **Medium models (100MB–1GB)** | .h5, .pth, .bin | Best for DL weights |
| **Large models (>1GB)** | .safetensors, .pb, sharded files | Faster, memory-safe |