AgroAI Models with Data Cleaning, Feature Selection, Platform and Input Types

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model Name | Type | Learning | Typical Use Case | Popular Tools/Libraries | Quantitative Data | Image/Text Data | Data Cleaning Methods | Feature Selection Methods | Supported Platforms | Input Type |
| MobileNetV2 | CNN (Deep Learning) | Supervised | Plant disease detection | TensorFlow, TFLite | No | Image | Image resizing, normalization | PCA, filter methods (based on CNN layers) | Android, Edge Devices | Image |
| EfficientNet | CNN (Deep Learning) | Supervised | High-accuracy plant image classification | TensorFlow, TFLite | No | Image | Noise reduction, augmentation | CNN feature maps | Android, Web | Image |
| ResNet50 | CNN (Deep Learning) | Supervised | Deep plant disease classification | TensorFlow, PyTorch | No | Image | Image normalization, RGB conversion | Layer activations, Grad-CAM | Cloud, GPU Servers | Image |
| VGG16 | CNN (Deep Learning) | Supervised | Simple image classification | Keras, TensorFlow | No | Image | Resize, pixel scaling | CNN filters | Web, Mobile | Image |
| DistilBERT | Transformer (NLP) | Supervised | Symptom classification from text | HuggingFace Transformers | No | Text | Tokenization, stopword removal | Attention scores | Web, NLP Platforms | Text |
| BERT | Transformer (NLP) | Supervised | Multilingual symptom classification | HuggingFace Transformers | No | Text | Lemmatization, sentence segmentation | Embedding layer pruning | Web, Cloud APIs | Text |
| LSTM | Recurrent NN | Supervised | Text pattern analysis | TensorFlow, Keras | No | Text | Padding, removing noise | Sequence importance analysis | Mobile, Cloud | Text |
| Naive Bayes | Probabilistic Model | Supervised | Basic text classification | scikit-learn, NLTK | No | Text | Lowercasing, punctuation removal | Chi-square, Mutual Information | Desktop, Cloud | Text |
| Random Forest | Ensemble Learning | Supervised | Crop recommendation, soil analysis | scikit-learn | Yes | No | Outlier removal, imputation | Feature importance ranking | Desktop, Cloud | Tabular |
| XGBoost | Gradient Boosting | Supervised | Crop recommendation, yield prediction | XGBoost | Yes | No | Missing value handling, normalization | SHAP, gain-based selection | Cloud, Desktop | Tabular |
| MLP | Neural Network | Supervised | Simple numeric predictors | Keras, TensorFlow | Yes | No | Standardization, min-max scaling | Correlation matrix, embedded selection | Mobile, Desktop | Tabular |
| Linear Regression | Statistical Model | Supervised | Predicting continuous values | scikit-learn, Statsmodels | Yes | No | Removing outliers, scaling | Backward elimination, Lasso | Desktop, Web | Tabular |
| Logistic Regression | Classification | Supervised | Binary classification | scikit-learn, TensorFlow, PyTorch | Yes | No | Binary encoding, standardization | Recursive feature elimination (RFE) | Mobile, Desktop | Tabular |
| Decision Trees | Tree-based | Supervised | Customer segmentation | scikit-learn, XGBoost | Yes | Less frequent | Handling missing values, duplicates | Gini importance, entropy gain | Desktop | Tabular |
| K-Nearest Neighbors | Instance-based | Supervised | Pattern recognition, image classification | scikit-learn | Yes | Yes | Normalization, feature scaling | Distance-based feature ranking | Desktop, Mobile | Image, Tabular |