```
# Data Analytics Internship - Task 4: Customer Churn Prediction
# Machine Learning project requirements
# Minimum Python version: 3.8+

# Core Data Processing
pandas>=1.5.0               # Data manipulation and analysis
numpy>=1.21.0               # Numerical computing and array operations

# Machine Learning Core
scikit-learn>=1.1.0         # ML algorithms, preprocessing, and
evaluation
scipy>=1.9.0                # Statistical functions and optimization

# Data Visualization
matplotlib>=3.5.0           # Basic plotting and visualization
seaborn>=0.11.2            # Statistical data visualization
plotly>=5.10.0             # Interactive visualizations (optional)

# Model Persistence and Deployment
joblib>=1.2.0              # Efficient model serialization
pickle                     # Built-in Python serialization (backup)

# Advanced ML Tools (Optional but Recommended)
imbalanced-learn>=0.9.0    # Handling imbalanced datasets (SMOTE, etc.)
xgboost>=1.6.0             # Gradient boosting algorithm (optional)
lightgbm>=3.3.0            # Fast gradient boosting (optional)

# Model Interpretation and Explainability
shap>=0.41.0               # Model explainability (optional)

# Development Environment
jupyter>=1.0.0             # Interactive notebooks
ipython>=7.0.0             # Enhanced Python shell
notebook>=6.4.0            # Jupyter notebook interface
```

```
# Data Quality and Validation
missingno>=0.5.0           # Missing data visualization (optional)

# Progress Tracking
tqdm>=4.64.0               # Progress bars for long operations

# Hyperparameter Optimization (Optional Advanced)
# optuna>=3.0.0            # Advanced hyperparameter tuning (uncomment if needed)
# hyperopt>=0.2.7          # Bayesian optimization (uncomment if needed)

# Web Framework for Model Deployment (Optional)
# flask>=2.2.0             # Lightweight web framework (uncomment for API)
# streamlit>=1.20.0        # Interactive web apps (uncomment for dashboard)

# Data Export and Reporting
openpyxl>=3.0.0           # Excel file operations
reportlab>=3.6.0          # PDF report generation (optional)

# Utility Libraries
python-dateutil>=2.8.0    # Date parsing and manipulation
warnings                  # Built-in warning control

# Statistical Analysis (Optional)
statsmodels>=0.13.0       # Advanced statistical analysis (optional)

# Performance Optimization
pyarrow>=8.0.0            # Fast columnar data operations (optional)

# Core Requirements Summary (Minimum Installation):
# pip install pandas numpy scikit-learn matplotlib seaborn jupyter joblib
```

```
# Full Installation Command:
# pip install -r requirements.txt

# Development Installation (includes optional packages):
# pip install pandas numpy scikit-learn matplotlib seaborn jupyter joblib imbalanced-learn tqdm plotly shap

# Production Deployment Additional Requirements:
# flask gunicorn redis celery (for API deployment)
# docker (for containerization)
# mlflow (for ML lifecycle management)

# Notes:
# - All core ML functionality works with the minimum requirements
# - Optional packages enhance capabilities but are not essential
# - Choose packages based on specific project needs and computational resources
# - Some packages may require additional system dependencies (e.g., C++ compilers)
```