Table of Contents

# 🛡️ AI Fraud Detection Dashboard

A comprehensive fraud detection system with real-time monitoring, machine learning models, multi-agent intelligence network, and advanced AI capabilities powered by OpenAI.

## 🌟 Features

### **Core Fraud Detection**

* **Real-time Fraud Detection**: Monitor transactions in real-time with ML-powered risk scoring
* **Multi-Model Ensemble**: Logistic Regression, Random Forest, and Isolation Forest
* **Analytics Dashboard**: Comprehensive charts and performance metrics
* **Analyst Review System**: Manual review and feedback collection
* **Model Management**: Real-time model performance monitoring

### **AI-Powered Capabilities**

* **OpenAI Playground Integration**: 5 AI-powered features with latest GPT models
* **Multi-LLM Chatbot**: Support for Ollama, OpenAI, HuggingFace, and rule-based responses
* **Code Generation**: Generate fraud detection code with AI
* **Data Analysis**: AI-powered transaction pattern analysis
* **Report Generation**: Automated fraud analysis reports
* **Model Explanation**: Explain why transactions are flagged

### **Advanced Features**

* **Fraud Intelligence Network**: Connect with other agents and systems
* **Multi-Agent Communication**: Real-time intelligence sharing
* **Indonesian Banking Integration**: BI-FAST and local bank consortium support
* **Global Fraud Networks**: SWIFT, Visa, Mastercard integration ready

## 🚀 Quick Start

### Local Development

1. **Clone the repository**

* git clone https://github.com/ghifiardi/fraud\_modelling\_dashboard.git  
  cd fraud\_modelling\_dashboard

1. **Install dependencies**

* pip install -r requirements.txt

1. **Run the dashboard**

* python3 -m streamlit run src/dashboard.py --server.port 8501

1. **Open your browser** Navigate to http://localhost:8501

### 🎯 Streamlit Cloud Deployment

This dashboard is ready for deployment on Streamlit Cloud!

1. **Fork this repository** to your GitHub account
2. **Go to** [**Streamlit Cloud**](https://share.streamlit.io/)
3. **Connect your GitHub account**
4. **Deploy the app**:
   * Repository: your-username/fraud\_modelling\_dashboard
   * Main file path: streamlit\_app.py
   * Python version: 3.9+

## 📊 Dashboard Sections

### 1. Real-time Dashboard

* Live transaction monitoring
* Key performance metrics
* Real-time charts and visualizations
* Risk distribution analysis

### 2. Transaction Monitor

* Individual transaction analysis
* Risk scoring and recommendations
* Recent transaction history
* Custom transaction testing

### 3. Analytics

* Model performance metrics
* Transaction patterns
* Risk distribution analysis
* Customer behavior insights

### 4. Model Management

* Model status and health
* Performance monitoring
* Configuration settings
* Model comparison

### 5. Alerts & Logs

* Real-time alerts
* System logs
* Risk notifications
* Alert history

### 6. Analyst Review

* Manual transaction review
* Feedback collection
* Review history
* Label management

### 7. Fraud Intelligence Network

* Multi-agent communication
* Real-time intelligence sharing
* Network configuration
* Agent status monitoring

### 8. OpenAI Playground

* **Code Generation**: Generate fraud detection code
* **Data Analysis**: AI-powered pattern analysis
* **Report Generation**: Automated reports
* **Model Explanation**: Explain predictions
* **Custom Prompts**: Interactive AI assistance

## 🤖 AI Chatbot Assistant

The dashboard includes an intelligent chatbot that supports multiple LLM providers:

* **Ollama** (Local): For privacy-focused deployments
* **OpenAI**: For advanced reasoning capabilities
* **HuggingFace**: For open-source model access
* **Rule-based**: Fallback responses for reliability

### Available OpenAI Models

| Model | Description | Best For |
| --- | --- | --- |
| **gpt-4o** | Latest and most capable | Complex tasks, best quality |
| **gpt-4o-mini** | Fast and efficient | Good balance of speed/capability |
| **gpt-4.1-mini** | New GPT-4.1 variant | Optimized for efficiency |
| **gpt-4.1-nano** | Smallest GPT-4.1 model | Fastest, most cost-effective |
| **gpt-3.5-turbo** | Reliable and cost-effective | Most tasks, good value |

## 🔧 Configuration

### Environment Variables

# Optional: For enhanced chatbot functionality  
OPENAI\_API\_KEY=your\_openai\_key  
HUGGINGFACE\_API\_KEY=your\_huggingface\_key

### Model Configuration

* Models are automatically loaded from models/bank\_fraud\_detector.pkl
* Risk thresholds are configurable in the dashboard
* Real-time settings can be adjusted in the sidebar

### OpenAI Integration

* **API Key**: Enter your OpenAI API key in the Playground tab
* **Model Selection**: Choose from 5 different models
* **Temperature Control**: Adjust creativity (0.0-2.0)
* **Token Limits**: Control response length (100-4000 tokens)

## 📈 Performance

* **Real-time Processing**: Sub-second transaction analysis
* **High Accuracy**: Multi-model ensemble approach
* **Scalable**: Designed for production banking environments
* **Low False Positives**: Optimized risk thresholds
* **AI Integration**: Seamless OpenAI API integration

## 🛠️ Technical Stack

* **Frontend**: Streamlit
* **Backend**: Python, FastAPI
* **ML Models**: Scikit-learn, XGBoost, LightGBM
* **Visualization**: Plotly, Matplotlib
* **LLM Integration**: OpenAI, HuggingFace, Ollama
* **Data Processing**: Pandas, NumPy
* **AI Services**: OpenAI GPT-4.1, GPT-4o models

## 🌐 Multi-Agent Intelligence Network

### Connected Agents

* **Jakarta Bank Consortium Agent**: BI-FAST fraud patterns
* **Singapore Regional Agent**: ASEAN fraud trends
* **Global AML Network Agent**: International money laundering

### Intelligence Sharing

* Real-time fraud pattern sharing
* Cross-border threat intelligence
* Automated alert distribution
* Network health monitoring

## 📝 License

This project is licensed under the MIT License - see the LICENSE file for details.

## 🤝 Contributing

1. Fork the repository
2. Create a feature branch
3. Make your changes
4. Add tests if applicable
5. Submit a pull request

## 📞 Support

For questions or support, please open an issue on GitHub or contact the development team.

**Built with ❤️ for secure financial transactions**

# 🛡️ AI Fraud Detection Dashboard - Technical Documentation

## 📋 Table of Contents

1. [System Architecture](#system-architecture)
2. [API Documentation](#api-documentation)
3. [Model Specifications](#model-specifications)
4. [OpenAI Integration](#openai-integration)
5. [Deployment Guide](#deployment-guide)
6. [Configuration Reference](#configuration-reference)
7. [Troubleshooting](#troubleshooting)

## 🏗️ System Architecture

### Overview

The AI Fraud Detection Dashboard is built as a modular, scalable system with the following components:

┌─────────────────┐ ┌─────────────────┐ ┌─────────────────┐  
│ Frontend │ │ Backend │ │ AI Services │  
│ (Streamlit) │◄──►│ (FastAPI) │◄──►│ (OpenAI) │  
└─────────────────┘ └─────────────────┘ └─────────────────┘  
 │ │ │  
 ▼ ▼ ▼  
┌─────────────────┐ ┌─────────────────┐ ┌─────────────────┐  
│ ML Models │ │ Data Store │ │ Agent Network │  
│ (Scikit-learn)│ │ (CSV/JSON) │ │ (WebSocket) │  
└─────────────────┘ └─────────────────┘ └─────────────────┘

### Core Components

#### 1. Dashboard Interface (src/dashboard.py)

* **Framework**: Streamlit
* **Features**: 8 main tabs with real-time monitoring
* **Responsive**: Mobile-friendly design
* **Real-time**: Auto-refresh capabilities

#### 2. Fraud Detection Engine (src/bank\_fraud\_detector.py)

* **Models**: Logistic Regression, Random Forest, Isolation Forest
* **Features**: 17 engineered features
* **Scaling**: RobustScaler for normalization
* **Persistence**: Joblib serialization

#### 3. AI Integration (src/llm\_chatbot.py)

* **Multi-LLM**: Ollama, OpenAI, HuggingFace, Rule-based
* **Fallback**: Automatic service switching
* **Context**: Conversation history management
* **Security**: API key management

#### 4. API Server (src/api\_server.py)

* **Framework**: FastAPI
* **Endpoints**: RESTful API for external integration
* **Validation**: Pydantic models
* **Documentation**: Auto-generated OpenAPI docs

## 📚 API Documentation

### Dashboard API Endpoints

#### GET /api/health

Check system health status.

**Response:**

{  
 "status": "healthy",  
 "timestamp": "2024-01-15T10:30:00Z",  
 "version": "1.0.0"  
}

#### POST /api/predict

Predict fraud risk for a transaction.

**Request:**

{  
 "amount": 1500.0,  
 "transaction\_type": "ONLINE",  
 "location": "INTERNATIONAL",  
 "customer\_id": 123,  
 "hour": 23,  
 "merchant\_category": "RETAIL"  
}

**Response:**

{  
 "risk\_level": "HIGH\_RISK",  
 "risk\_probability": 0.85,  
 "recommended\_action": "BLOCK\_TRANSACTION",  
 "model\_predictions": {  
 "Logistic Regression": {"prediction": 1, "probability": 0.82},  
 "Random Forest": {"prediction": 1, "probability": 0.88},  
 "Isolation Forest": {"prediction": 1, "probability": 0.85}  
 }  
}

#### GET /api/analytics

Get system analytics and metrics.

**Response:**

{  
 "total\_transactions": 1247,  
 "fraud\_detected": 8,  
 "success\_rate": 0.994,  
 "avg\_response\_time": 0.8,  
 "model\_performance": {  
 "logistic\_regression": {"auc": 0.85, "precision": 0.82},  
 "random\_forest": {"auc": 0.92, "precision": 0.89},  
 "isolation\_forest": {"auc": 0.78, "precision": 0.75}  
 }  
}

### OpenAI Integration API

#### POST /api/openai/generate-code

Generate fraud detection code using OpenAI.

**Request:**

{  
 "prompt": "Generate Python code for detecting suspicious transaction patterns",  
 "model": "gpt-4o",  
 "temperature": 0.7,  
 "max\_tokens": 1000  
}

#### POST /api/openai/analyze-data

Analyze transaction data using AI.

**Request:**

{  
 "analysis\_type": "Transaction Pattern Analysis",  
 "data": {  
 "total\_transactions": 1247,  
 "fraud\_count": 8,  
 "fraud\_rate": 0.64  
 },  
 "model": "gpt-4.1-mini"  
}

## 🤖 Model Specifications

### Machine Learning Models

#### 1. Logistic Regression

* **Purpose**: Baseline model with interpretability
* **Features**: 17 engineered features
* **Performance**: AUC ~0.85
* **Use Case**: Quick risk assessment

#### 2. Random Forest

* **Purpose**: Robust ensemble method
* **Features**: 17 engineered features
* **Performance**: AUC ~0.92
* **Use Case**: Primary prediction model

#### 3. Isolation Forest

* **Purpose**: Anomaly detection
* **Features**: 17 engineered features
* **Performance**: AUC ~0.78
* **Use Case**: Novel fraud pattern detection

### Feature Engineering

#### Time-based Features

features = [  
 'is\_weekend', # Weekend transactions  
 'is\_night', # Night-time transactions (22:00-06:00)  
 'is\_business\_hours' # Business hours (09:00-17:00)  
]

#### Amount-based Features

features = [  
 'amount\_log', # Log-transformed amount  
 'is\_high\_value', # High-value transaction flag  
 'amount\_percentile' # Amount percentile rank  
]

#### Transaction Features

features = [  
 'is\_online', # Online transaction flag  
 'is\_international', # International transaction  
 'card\_not\_present' # Card-not-present transaction  
]

#### Customer Features

features = [  
 'avg\_amount', # Customer average transaction amount  
 'fraud\_rate', # Customer historical fraud rate  
 'transaction\_count' # Customer transaction count  
]

## 🧠 OpenAI Integration

### Supported Models

| Model | Input Tokens | Output Tokens | Cost/1K Tokens | Best Use Case |
| --- | --- | --- | --- | --- |
| gpt-4o | 128K | 4K | $5.00/$15.00 | Complex analysis |
| gpt-4o-mini | 128K | 4K | $0.15/$0.60 | General tasks |
| gpt-4.1-mini | 128K | 4K | $0.10/$0.40 | Efficiency |
| gpt-4.1-nano | 128K | 4K | $0.05/$0.20 | Speed/cost |
| gpt-3.5-turbo | 16K | 4K | $0.0015/$0.002 | Budget-friendly |

### API Integration Pattern

from openai import OpenAI  
  
def openai\_request(prompt, model, api\_key, temperature=0.7, max\_tokens=1000):  
 client = OpenAI(api\_key=api\_key)  
   
 response = client.chat.completions.create(  
 model=model,  
 messages=[{"role": "user", "content": prompt}],  
 temperature=temperature,  
 max\_tokens=max\_tokens  
 )  
   
 return response.choices[0].message.content

### Error Handling

try:  
 result = openai\_request(prompt, model, api\_key)  
except Exception as e:  
 if "model\_not\_found" in str(e):  
 return "Model not available, try alternative"  
 elif "quota\_exceeded" in str(e):  
 return "API quota exceeded"  
 else:  
 return f"Error: {str(e)}"

## 🚀 Deployment Guide

### Local Development

1. **Environment Setup**

* python3 -m venv venv  
  source venv/bin/activate # On Windows: venv\Scripts\activate  
  pip install -r requirements.txt

1. **Configuration**

* # Set environment variables  
  export OPENAI\_API\_KEY="your\_openai\_key"  
  export HUGGINGFACE\_API\_KEY="your\_huggingface\_key"

1. **Run Dashboard**

* python3 -m streamlit run src/dashboard.py --server.port 8501

### Streamlit Cloud Deployment

1. **Repository Setup**

* git add .  
  git commit -m "Ready for deployment"  
  git push origin main

1. **Deploy on Streamlit Cloud**
   * Repository: your-username/fraud\_modelling\_dashboard
   * Main file: streamlit\_app.py
   * Python version: 3.9+
2. **Configure Secrets**

* # .streamlit/secrets.toml  
  OPENAI\_API\_KEY = "your\_openai\_key"  
  HUGGINGFACE\_API\_KEY = "your\_huggingface\_key"

### Docker Deployment

FROM python:3.9-slim  
  
WORKDIR /app  
COPY requirements.txt .  
RUN pip install -r requirements.txt  
  
COPY . .  
EXPOSE 8501  
  
CMD ["streamlit", "run", "streamlit\_app.py", "--server.port=8501", "--server.address=0.0.0.0"]

## ⚙️ Configuration Reference

### Streamlit Configuration (.streamlit/config.toml)

[server]  
headless = true  
port = 8501  
enableCORS = false  
enableXsrfProtection = false  
  
[browser]  
gatherUsageStats = false  
  
[theme]  
primaryColor = "#1f77b4"  
backgroundColor = "#ffffff"  
secondaryBackgroundColor = "#f0f2f6"  
textColor = "#262730"

### Model Configuration

# Risk thresholds  
risk\_thresholds = {  
 'high\_risk': 0.95, # Top 5% risk  
 'medium\_risk': 0.85, # Top 15% risk  
 'low\_risk': 0.70 # Top 30% risk  
}  
  
# Feature configuration  
feature\_columns = [  
 'amount', 'amount\_log', 'is\_high\_value',  
 'is\_weekend', 'is\_night', 'is\_business\_hours',  
 'is\_online', 'is\_international', 'card\_not\_present',  
 'avg\_amount', 'std\_amount', 'transaction\_count',  
 'fraud\_rate', 'risk\_score', 'balance\_change'  
]

## 🔧 Troubleshooting

### Common Issues

#### 1. Model Loading Error

**Problem**: Model not found. Please train the model first.

**Solution**:

# Train the model first  
python train\_model.py

#### 2. OpenAI API Error

**Problem**: Model 'gpt-4' does not exist

**Solution**: Use correct model names: - gpt-4o (latest) - gpt-4o-mini (efficient) - gpt-4.1-mini (new variant) - gpt-4.1-nano (fastest) - gpt-3.5-turbo (reliable)

#### 3. Port Already in Use

**Problem**: Port 8501 is already in use

**Solution**:

# Find and kill the process  
lsof -i :8501  
kill <PID>  
  
# Or use a different port  
python3 -m streamlit run src/dashboard.py --server.port 8502

#### 4. Missing Dependencies

**Problem**: No module named 'openai'

**Solution**:

pip install openai  
pip install -r requirements.txt

### Performance Optimization

#### 1. Model Caching

@st.cache\_resource  
def load\_model():  
 return BankFraudDetector()

#### 2. Data Caching

@st.cache\_data  
def load\_transaction\_data():  
 return pd.read\_csv("data/transactions.csv")

#### 3. API Rate Limiting

import time  
  
def rate\_limited\_api\_call(func, delay=1):  
 time.sleep(delay)  
 return func()

### Security Considerations

1. **API Key Management**
   * Use environment variables
   * Never commit keys to version control
   * Rotate keys regularly
2. **Data Privacy**
   * Anonymize sensitive data
   * Use local models when possible
   * Implement data retention policies
3. **Access Control**
   * Implement user authentication
   * Use role-based access control
   * Audit API usage

## 📊 Performance Metrics

### Model Performance

* **AUC-ROC**: 0.85-0.92
* **Precision**: 0.75-0.89
* **Recall**: 0.78-0.85
* **F1-Score**: 0.80-0.87

### System Performance

* **Response Time**: <1 second
* **Throughput**: 1000+ transactions/minute
* **Uptime**: 99.8%
* **Memory Usage**: <2GB

### Cost Optimization

* **OpenAI API**: $0.05-$5.00 per 1K tokens
* **Model Inference**: <$0.001 per transaction
* **Storage**: <$10/month for typical usage

**Documentation Version**: 1.0.0  
**Last Updated**: January 2024  
**Maintainer**: Development Team