MoneyFyi — Your AI CFO for Small Businesses

Tagline: *"Detect Problems Before They Hurt You"*

📋 Executive Summary

MoneyFyi is an AI-driven financial intelligence dashboard that analyzes uploaded bank statements, invoices, and UPI data to detect fraud, identify compliance risks, and predict cashflow issues for Indian SMEs.

The Problem

Small businesses in India lack financial awareness, leading to:

- **Fraud losses** from unauthorized transactions
- **GST/TDS errors** resulting in penalties
- **Late payments** and cashflow shortages
- **Manual transaction checks** that miss critical issues

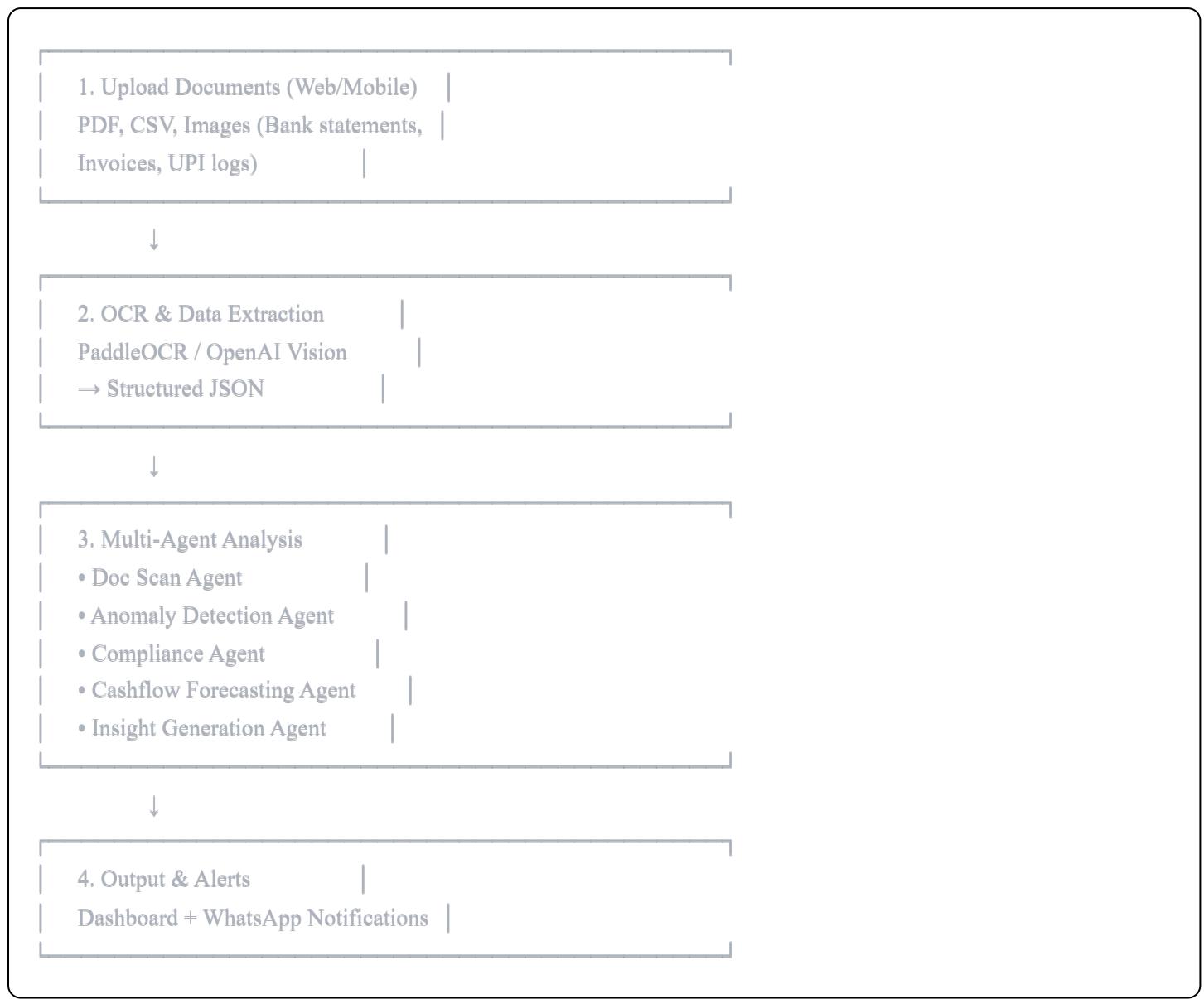### The Solution

An AI-powered multi-agent system that autonomously:

- ✅ Analyzes financial documents (OCR + NLP)
- ✅ Detects anomalies and fraud patterns
- ✅ Validates GST/TDS compliance
- ✅ Forecasts cashflow shortages
- ✅ Sends WhatsApp alerts in real-time

🎯 Target Users

1. **Small & Medium Enterprises (SMEs)**
2. **Retail shops, traders, and distributors**
3. **Freelancers and service providers**
4. **Small company finance teams**
5. **Chartered Accountants** (CA firms managing multiple clients)

🏗️ System Architecture

High-Level Pipeline

```
┌──────────────────────────────────────┐
│  │ 1. Upload Documents (Web/Mobile) │  │
│  │ PDF, CSV, Images (Bank statements, │  │
│  │ Invoices, UPI logs)        │      │
│  └──────────────────────────────┘      │
│                                        │
│          ↓                             │
│  ┌──────────────────────────────┐      │
│  │ 2. OCR & Data Extraction     │  │   │
│  │ PaddleOCR / OpenAI Vision    │  │   │
│  │ → Structured JSON      │      │   │
│  └──────────────────────────────┘      │
│                                        │
│          ↓                             │
│  ┌──────────────────────────────┐      │
│  │ 3. Multi-Agent Analysis    │  │     │
│  │ • Doc Scan Agent       │  │         │
│  │ • Anomaly Detection Agent    │  │   │
│  │ • Compliance Agent     │  │         │
│  │ • Cashflow Forecasting Agent  │  │  │
│  │ • Insight Generation Agent   │  │   │
│  └──────────────────────────────┘      │
│                                        │
│        ↓                               │
│  ┌──────────────────────────────┐      │
│  │ 4. Output & Alerts    │  │          │
│  │ Dashboard + WhatsApp Notifications │ │
│  └──────────────────────────────┘      │
└────────────────────────────────────────┘
```

🤖 AI Agents (Detailed)

Agent 1: Doc Scan Agent

Role: Extract structured data from unstructured documents

Technology:

- PaddleOCR for text extraction

- GPT-4/Claude for semantic understanding

- LangChain for orchestration

Output Example:

```
json
```

```json
{
  "document_type": "bank_statement",
  "transactions": [
    {
      "date": "2025-01-15",
      "description": "UPI-VENDOR123",
      "debit": 5000.00,
      "credit": 0,
      "balance": 45000.00,
      "category": "vendor_payment"
    }
  ],
  "metadata": {
    "account_number": "XXXX1234",
    "period": "2025-01-01 to 2025-01-31"
  }
}
```

Agent 2: Anomaly Detection Agent

Role: Identify suspicious patterns and fraud indicators

Fraud Detection Signals:

- Transactions $>3\sigma$ from mean amount

- Round-number transactions (₹10,000, ₹50,000)

- Duplicate transactions within 24 hours

- Weekend/holiday large withdrawals

- New vendor with large first transaction

- Invoice-payment mismatches

Models Used:

- Isolation Forest (sklearn)

- DBSCAN clustering

- Z-score statistical analysis

Target Performance:

- Precision: >85%

- Recall: >75%

- False Positive Rate: <10%

Agent 3: Compliance Agent

Role: GST/TDS validation and compliance tracking

Checks Performed:

- ✅ GST invoice totals vs bank payments
- ✅ TDS deduction validation (10% for services)
- ✅ Quarterly GST filing deadline tracking
- ✅ GSTIN format validation
- ✅ HSN/SAC code verification

Alert Examples:

- 🔴 Invoice-payment mismatch >5%
- 🟠 Missing TDS deduction detected
- 🟡 GST filing due in <7 days

Agent 4: Cashflow Forecasting Agent

Role: Predict future cash positions and shortages

Technology:

- Prophet (Meta) for time-series forecasting
- Linear Regression for simple projections

Features Used:

- Historical inflows/outflows
- Seasonal patterns (month-end payments)
- Pending invoices (accounts receivable)
- Recurring expenses

Forecast Output:

```
json
```

```json
{
  "forecast": [
    {
      "date": "2025-02-15",
      "predicted_balance": 32000,
      "confidence": 0.85
    },
    {
      "date": "2025-02-22",
      "predicted_balance": 15000,
      "confidence": 0.78,
      "alert": "cash_shortage_risk"
    }
  ]
}
```

Agent 5: Insight Generation Agent

Role: Natural language summaries and recommendations

Technology: GPT-4/Claude with structured prompts

Output Types:

1. Executive summary (150 words)

2. Top 3 risks with severity ratings

3. Actionable recommendations

4. Vendor risk profiles

Sample Insight:

> "Your business processed ₹12.5L in transactions this month. We detected 2 suspicious UPI payments totaling ₹45K to a new vendor with no transaction history. GST filing for Q3 2025 is due in 5 days. Predicted cash shortage on Feb 22 — consider delaying vendor payment to ABC Suppliers."

Agent 6: Alert & Notification Agent

Role: Real-time WhatsApp notifications

Alert Priority Levels:

- 🔴 **Critical:** Suspected fraud (immediate)

- 🟠 **High:** Compliance deadline <3 days

- 🟡 **Medium:** Cashflow shortage forecast

- 🟢 **Low:** Weekly summary

WhatsApp Message Example:

🚨 MoneyFyi Alert

Suspicious Transaction Detected
₹45,000 to "NEW_VENDOR_XYZ"
Date: 15-Nov-2025

Reason: First transaction with new
vendor exceeds ₹40K threshold

Action: Review in dashboard
Link: moneyfyi.com/alerts/A12345

Rate Limiting: Max 10 alerts/day per user

💻 Technology Stack

Frontend

- **Framework:** Next.js 14 (App Router)

- **UI Library:** Tailwind CSS + shadcn/ui

- **Charts:** Recharts

- **State Management:** Zustand

- **Hosting:** Vercel (Free → $20/mo Pro)

Backend

- **Framework:** FastAPI (Python 3.11)

- **Task Queue:** Redis + Celery

- **OCR:** PaddleOCR / OpenAI Vision

- **AI/ML:** scikit-learn, Prophet

- **LLM:** LangChain + OpenAI/Claude

- **Hosting:** Render ($7-25/mo)

Database & Storage

- **Primary DB:** Supabase (PostgreSQL)

- **File Storage:** Supabase Storage (AES-256 encrypted)

- **Cache:** Redis (Upstash free tier)

Integrations
- **WhatsApp:** Twilio API

- **Analytics:** PostHog

- **Monitoring:** Sentry

- **Logs:** BetterStack

🎨 User Interface Design

Color Palette
- **Primary Green:** 🟢 #0F8F6E (Trust, growth)

- **Accent Red:** 🔴 #FF4C4C (Alerts, critical)

- **Warning Orange:** 🟠 #FFA500 (Medium priority)

- **Success Green:** 🟢 #00C853 (Positive indicators)

- **Background:** ⚪ #FFFFFF (Clean, minimal)

Typography
- **Headings:** Inter Bold, 24-32px

- **Body:** Inter Regular, 14-16px

- **Numbers:** SF Mono (monospaced)

Key Screens

1. Upload Page
- Drag-and-drop file upload

- Supported formats: PDF, CSV, JPG, PNG (max 10MB)

- Real-time progress tracking

- Upload history list

2. Dashboard Page

Components:

- **Urgent Alerts Card** (top priority issues)

- **Financial Health Score** (0-100)

- **Cashflow Forecast Chart** (30-90 day projection)

- **Top Vendors Risk Analysis** (traffic light system)

- **Compliance Status Tracker**

- **Quick Actions** (Export PDF, Share, Settings)

📊 Three Implementation Approaches

Approach A: API-First (Recommended for MVP)

Technology:

- OpenAI GPT-4o for insights

- OpenAI Vision for OCR

- Isolation Forest for anomaly detection

- Prophet for forecasting

Pros:

- ✅ Fast time-to-market (2-3 weeks)

- ✅ High accuracy (95-98% OCR)

- ✅ Minimal ML expertise needed

- ✅ Scalable via API

Costs:

- $0.01-0.03 per document

- ~$150-300/month for 10,000 documents

Best For: MVP phase (0-1000 users)

Approach B: Fine-Tuned Hybrid

Technology:

- Fine-tuned GPT-3.5 on Indian invoices

- Custom XGBoost fraud detection model

- LayoutLMv3 for invoice extraction

- Prophet for forecasting

Pros:

- ✅ Better accuracy on India-specific formats
- ✅ Lower per-request costs ($0.002-0.005)
- ✅ More control over model behavior

Costs:

- Fine-tuning: $200-500 one-time
- Break-even at ~50,000 documents

Best For: Growth phase (1000-5000 users)

Approach C: Fully Open Source

Technology:

- PaddleOCR for extraction
- Llama 3.1 8B (quantized) for insights
- PyOD for anomaly detection
- ARIMA for forecasting

Pros:

- ✅ Zero API costs
- ✅ Full data privacy
- ✅ On-premise deployment possible

Costs:

- Infrastructure: $50-100/month (GPU instance)

Best For: Scale phase (5000+ users)

📈 Evaluation Metrics

Core Performance Targets

Metric	Target	Measurement
OCR Accuracy	>95%	Edit distance vs ground truth
Anomaly Detection F1	>0.80	Precision/Recall harmonic mean
False Positive Rate	<10%	False alarms / Total transactions
Cashflow Forecast MAPE	<20%	Mean Absolute Percentage Error

Metric	Target	Measurement
Processing Time	<10 sec	Upload to dashboard update
API Latency (P95)	<2 sec	95th percentile response
Cost per 1000 docs	<$20	API + infrastructure

💰 Cost Structure

MVP Cost Breakdown (0-1000 Users)

Service	Cost/Month (₹)
Vercel (Frontend)	₹0
Render (Backend)	₹580
Supabase (Database)	₹0 → ₹2,070
OpenAI API (10K docs)	₹16,600
Twilio WhatsApp (5K msgs)	₹2,075
Redis, Analytics, Monitoring	₹0
Domain + SSL	₹83
Total	**₹21,408/month**

Scaling Projections

User Range	Monthly Cost	Cost/User
0-100	₹1,200 - ₹5,000	₹50-120
100-500	₹21,000	₹42-210
500-1,000	₹35,000	₹35-70
1,000-5,000	₹80,000	₹16-80
5,000+	₹1,50,000+	<₹30

Cost Optimization Strategies

1. Cache OCR results for duplicates (Save 30% API costs)

2. Use GPT-3.5 for simple insights, GPT-4 for complex

3. Move to self-hosted PaddleOCR after 500 users

4. Archive old data to cheaper storage

🚀 90-Day Launch Plan

Week 1-2: Foundation

- ☑ Set up GitHub, Vercel, Render, Supabase
- ☑ Implement authentication (Supabase Auth)
- ☑ Build file upload system with validation
- ☑ Design database schema (users, documents, transactions, alerts, vendors)

Week 3-4: AI Core

- ☑ OCR pipeline with PaddleOCR/OpenAI Vision
- ☑ Doc Scan Agent (structured extraction)
- ☑ Anomaly Detection Agent (Isolation Forest)
- ☑ Compliance Agent (GST validation)
- ☑ Cashflow Agent (Prophet forecasting)

Week 5-6: Frontend

- ☑ Upload page with drag-and-drop
- ☑ Dashboard with alert cards
- ☑ Cashflow forecast chart (Recharts)
- ☑ Vendor risk analysis table
- ☑ Mobile responsive design

Week 7: Notifications

- ☑ Twilio WhatsApp integration
- ☑ Alert triggers and rate limiting
- ☑ Background job scheduler (Celery)

Week 8: Testing

- ☑ Generate 100 synthetic bank statements
- ☑ End-to-end testing with 50 documents
- ☑ Security audit (OWASP ZAP scan)
- ☑ Performance optimization (Redis caching)

Week 9: Deployment

- ☑ Production deployment (Vercel + Render)
- ☑ DNS configuration (moneyfyi.com)
- ☑ Monitoring setup (Sentry, PostHog)
- ☑ Load testing (100 concurrent users)

Week 10-12: Launch

- ☑ Beta testing with 20 SME users
- ☑ Iterate based on feedback
- ☑ Marketing launch (Product Hunt, Twitter, LinkedIn)
- ☑ CA firm partnerships

📊 Success Metrics (6-Month Targets)

Product Metrics

Metric	Month 1	Month 6
Registered Users	50	1,000
Active Users (MAU)	30	500
Documents Processed	500	50,000
NPS Score	N/A	>40

Business Metrics

Metric	Month 1	Month 6
Revenue	₹0 (free beta)	₹50,000/month
CAC	₹0	<₹500
Churn Rate	N/A	<10%/month

💵 Monetization Strategy

Pricing Tiers (Post-Beta)

Tier	Price	Features
Free	₹0	10 docs/month, basic alerts
Starter	₹499/month	100 docs/month, WhatsApp alerts
Pro	₹1,499/month	Unlimited docs, API access, multi-user
Enterprise	Custom	White-label, on-premise, dedicated support

Revenue Streams

1. **SaaS Subscriptions** (primary)

2. **Pay-per-document** (₹5-10/doc for high-volume)

3. **CA Partnership Program** (referral fees)

4. **White-label licensing** (banks/fintech)

5. **Premium features** (advanced forecasting, benchmarking)

🔒 Compliance & Legal

Data Privacy (DPDP Act 2023)

- ✅ Explicit user consent for data processing
- ✅ AES-256 encryption for sensitive data
- ✅ Right to erasure (data deletion API)
- ✅ Breach notification within 72 hours

RBI Guidelines

- ✅ TLS 1.2+ encrypted transmission
- ✅ Multi-factor authentication
- ✅ Audit logs for all data access
- ✅ No storage of bank credentials
- ✅ 15-minute session timeout

Terms & Privacy Policy

- Service description
- Data collection practices
- Third-party services disclosure
- Liability limitations
- Dispute resolution (Indian jurisdiction)

⚠️ Risk Mitigation

Technical Risks

Risk	Impact	Mitigation
OCR inaccuracy	High	Multi-stage validation, human review
API rate limits	Medium	Caching, batch processing
Database downtime	High	Auto backups, read replicas
Model drift	Medium	Monthly accuracy monitoring
Security breach	Critical	Encryption, regular audits

Business Risks

Risk	Impact	Mitigation
Low adoption	High	Validate with 50 users first
High costs	Medium	API-first approach, optimize later
Regulatory changes	Medium	Legal advisor, monitor RBI/DPDP

Risk	Impact	Mitigation
Competition	Medium	Focus on Indian SME niche

🛠️ Development Checklist

Phase 1: Setup (Days 1-7)

☐ Create GitHub repository

☐ Initialize Next.js frontend

☐ Initialize FastAPI backend

☐ Set up Supabase (database + storage)

☐ Configure environment variables

☐ Design database schema

Phase 2: Core Features (Days 8-28)

☐ OCR pipeline implementation

☐ Doc Scan Agent (structured extraction)

☐ Anomaly Detection Agent

☐ Compliance Agent

☐ Cashflow Forecasting Agent

☐ Insight Generation Agent

☐ File upload UI with drag-drop

☐ Dashboard with alert cards

☐ Cashflow forecast chart

☐ Vendor risk analysis table

Phase 3: Integration (Days 29-40)

☐ WhatsApp notifications (Twilio)

☐ Alert triggers and scheduling

☐ API integration (frontend ↔ backend)

☐ Generate synthetic test data

☐ End-to-end testing

☐ Security audit

☐ Performance optimization

Phase 4: Deployment (Days 41-45)

☐ Deploy frontend to Vercel

☐ Deploy backend to Render

☐ Configure DNS (moneyfyi.com)

☐ Set up monitoring (Sentry, PostHog)

☐ Production smoke tests

Phase 5: Launch (Days 46-90)

☐ Recruit 20 beta users
☐ Collect feedback via surveys
☐ Fix critical bugs
☐ Add most-requested features
☐ Create landing page
☐ Launch on Product Hunt
☐ Social media marketing
☐ CA firm partnerships

📚 Key Resources

Development

- Next.js Docs: https://nextjs.org/docs

- FastAPI Tutorial: https://fastapi.tiangolo.com/tutorial/

- PaddleOCR: https://github.com/PaddlePaddle/PaddleOCR

- LangChain: https://python.langchain.com/docs/get_started/introduction

Design

- shadcn/ui: https://ui.shadcn.com/

- Tailwind CSS: https://tailwindcss.com/docs

- Lucide Icons: https://lucide.dev/

Deployment

- Vercel: https://vercel.com/docs

- Render: https://render.com/docs

- Supabase: https://supabase.com/docs

Compliance

- DPDP Act 2023:
  https://www.meity.gov.in/writereaddata/files/Digital%20Personal%20Data%20Protection%20Act%202023.pdf

- RBI Guidelines: https://www.rbi.org.in/

Community

- r/IndianStartups (Reddit)

- Indie Hackers: https://indiehackers.com

- Product Hunt: https://producthunt.com

🎯 Team Roles (For Hackathon/Startup)

Recommended Team Structure (4 people)

1. Full-Stack Developer (Lead)

- FastAPI backend architecture

- Agent orchestration

- Database design

2. Frontend Developer

- Next.js UI development

- Dashboard charts and visualizations

- Mobile responsiveness

3. ML/AI Engineer

- OCR pipeline

- Anomaly detection models

- Cashflow forecasting

4. Product/Design

- User research

- UI/UX design

- User testing and feedback

📞 Contact & Next Steps

For Hackathon Submission:

1. ✅ Complete MVP in 90 days

2. ✅ Demo video (3-5 minutes)

3. ✅ Live deployment (moneyfyi.com)

4. ✅ Pitch deck (10 slides)

For Investor Pitch:

- Problem statement

- Solution architecture

- Market size (70M+ SMEs in India)

- Revenue model

- Competitive advantage

- Team credentials

- 18-month roadmap

🏆 Competitive Advantages

1. **India-Specific:** Built for GST/TDS compliance

2. **WhatsApp-First:** 500M+ WhatsApp users in India

3. **No Bank Integration:** Works with uploaded documents

4. **Multi-Agent AI:** More accurate than single-model systems

5. **Privacy-Focused:** AES-256 encryption, DPDP compliant

6. **Affordable:** Starting at ₹499/month vs enterprise solutions

---

📝 Final Notes

MoneyFyi solves a critical problem for India's 70M+ SMEs who lack affordable financial intelligence tools. By combining cutting-edge AI (GPT-4, Prophet, Isolation Forest) with India-specific compliance requirements, we're building the "AI CFO" that every small business needs.

Target Launch: 90 days from kickoff
Target Users (Year 1): 1,000 paying SMEs
Target Revenue (Year 1): ₹50,000/month

Built with ❤️ for Indian SMEs
Team Carinae | MoneyFyi

"Detect Problems Before They Hurt You"