**1️⃣ User Profile & Static Inputs (One-Time Inputs)**

These define the user's financial background and constraints.

| **Feature** | **Description** | **Type** |
| --- | --- | --- |
| **Monthly Income** | Salary, business earnings, passive income, etc. | Numeric |
| **Fixed Expenses** | Rent, EMI, Insurance, Subscriptions, Loan Payments | Numeric |
| **Financial Goals** | E.g., Save ₹10,000/month, Buy a house in 3 years | Categorical / Numeric |
| **Spending Limit Preferences** | User-defined budget rules (e.g., max 10% for entertainment) | Numeric |
| **Credit Card Debt** | Outstanding balances, monthly dues | Numeric |

**2️⃣ Real-Time Transaction Data (Fetched via API from AA)**

This data is **fetched regularly** from the user's **bank transactions** to train & update the AI model.

| **Feature** | **Description** | **Type** |
| --- | --- | --- |
| **Transaction Amount** | Value of each transaction | Numeric |
| **Transaction Date & Time** | Time when the transaction occurred | Timestamp |
| **Merchant Name / Description** | "Amazon," "Swiggy," "HDFC EMI," etc. | Text |
| **Transaction Category** | (Food, Rent, EMI, Travel, Shopping, etc.) | Categorical |
| **Payment Mode** | UPI, Debit Card, Credit Card, Wallet | Categorical |
| **Account Type** | Savings, Credit, Wallet, Loan | Categorical |
| **Recurring Transaction?** | Yes/No (Identifies rent, EMI, subscriptions) | Boolean |

**🔹 Data Source:**

* **Bank APIs via Account Aggregators (Finvu, INK AA, etc.).**
* **UPI APIs for tracking live spending (PhonePe, Google Pay).**
* **Credit Card APIs (if integrated).**

**3️⃣ Derived Features (AI-Generated)**

These features are calculated based on historical data and user behavior.

| **Feature** | **Description** | **Type** |
| --- | --- | --- |
| **Average Monthly Spend** | Mean of last 3-6 months' transactions | Numeric |
| **Expense Variability Score** | Standard deviation in spending patterns | Numeric |
| **Fixed vs. Variable Expense Ratio** | % of spending that is fixed vs. discretionary | Numeric |
| **Savings Rate** | % of income saved per month | Numeric |
| **Upcoming Large Payments** | AI-detected upcoming EMI, insurance, annual fees | Boolean |
| **Over Budget Alerts** | Flag if spending crosses 80% of the limit | Boolean |

**4️⃣ Historical Spending Patterns (For Forecasting)**

To predict future expenses, we need a **time-series dataset**.

| **Feature** | **Description** | **Type** |
| --- | --- | --- |
| **Past 6-12 Months Transactions** | Time-stamped spending history | Time-Series |
| **Seasonal Spending Patterns** | E.g., Higher expenses in festive months | Time-Series |
| **Unexpected Spikes in Spending** | AI-detected anomaly patterns | Numeric |
| **Savings Growth Over Time** | Trend of savings accumulation | Numeric |

🔹 **Used for:**

* Training the **LSTM model** for **expense forecasting**.
* Detecting **anomalies & overspending trends**.

**1️⃣ Sample Dataset: Bank Transactions (Categorization & Tracking)**

This dataset contains **bank transactions** with labels for different spending categories.

| **Transaction ID** | **Date** | **Amount (₹)** | **Merchant / Description** | **Category** | **Payment Mode** |
| --- | --- | --- | --- | --- | --- |
| 1001 | 2024-01-01 | 12,500 | HDFC Bank - EMI | EMI & Loans | Auto-Debit |
| 1002 | 2024-01-02 | 200 | Swiggy | Food & Dining | UPI |
| 1003 | 2024-01-03 | 8,000 | Amazon | Shopping | Credit Card |
| 1004 | 2024-01-05 | 15,000 | Rent Payment | Fixed Expense | Bank Transfer |
| 1005 | 2024-01-06 | 500 | Ola Cabs | Travel | Debit Card |
| 1006 | 2024-01-08 | 1,000 | Netflix Subscription | Entertainment | Auto-Debit |

🔹 **Use Case**: Train an NLP or ML model to classify transactions into predefined categories.

**2️⃣ Sample Dataset: Monthly Budget Summary (User Profiling & Budgeting Strategies)**

This dataset helps in **analyzing spending behavior** and **choosing the best budget strategy**.

| **User ID** | **Income (₹)** | **Fixed Expenses (₹)** | **Variable Expenses (₹)** | **Savings (₹)** | **Preferred Budget Plan** |
| --- | --- | --- | --- | --- | --- |
| U001 | 80,000 | 30,000 | 25,000 | 25,000 | 50-30-20 Budgeting |
| U002 | 50,000 | 20,000 | 20,000 | 10,000 | Zero-Based Budgeting |
| U003 | 1,20,000 | 40,000 | 50,000 | 30,000 | Envelope Budgeting |
| U004 | 60,000 | 25,000 | 25,000 | 10,000 | 50-30-20 Budgeting |

🔹 **Use Case**: Identify spending patterns and suggest personalized **budget strategies**.

**3️⃣ Sample Dataset: Expense Forecasting (Time-Series Data for LSTM Model)**

This dataset helps in **predicting future expenses** using historical spending patterns.

| **Month** | **Salary (₹)** | **Fixed Expenses (₹)** | **Variable Expenses (₹)** | **Total Spent (₹)** | **Savings (₹)** |
| --- | --- | --- | --- | --- | --- |
| Jan 24 | 80,000 | 30,000 | 25,000 | 55,000 | 25,000 |
| Feb 24 | 80,000 | 30,000 | 28,000 | 58,000 | 22,000 |
| Mar 24 | 80,000 | 30,000 | 26,000 | 56,000 | 24,000 |
| Apr 24 | 80,000 | 30,000 | 29,000 | 59,000 | 21,000 |
| May 24 | 80,000 | 30,000 | 27,000 | 57,000 | 23,000 |

🔹 **Use Case**: Train an **LSTM model** to predict next month's expenses and suggest savings plans.

**4️⃣ Sample Dataset: Anomaly Detection (Unusual Spending Alerts)**

This dataset flags **suspicious or unexpected transactions**.

| **Transaction ID** | **Date** | **Amount (₹)** | **Merchant** | **Category** | **Anomaly?** |
| --- | --- | --- | --- | --- | --- |
| 1001 | 2024-01-01 | 50,000 | Jewelry Store | Shopping | Yes |
| 1002 | 2024-01-02 | 200 | Swiggy | Food | No |
| 1003 | 2024-01-03 | 8,000 | Amazon | Shopping | No |
| 1004 | 2024-01-04 | 60,000 | Car Repair | Maintenance | Yes |
| 1005 | 2024-01-05 | 500 | Ola Cabs | Travel | No |

🔹 **Use Case**: Use **Isolation Forest or Autoencoders** to detect **unexpected spending behavior**.