

**AI -Powered Personalized Recommendation system - GolDEN Basket**

Customer’s Personalized wealth management recommendation system



March 26, 2025

Team – Operation AI

Table of Contents

[**Summary 2**](#_Toc193660623)

[**Who is this for? 2**](#_Toc193660624)

[**Problem Statement – Why Change is Needed 3**](#_Toc193660625)

[**Current Market Landscape 3**](#_Toc193660626)

[**Our Solution : 4**](#_Toc193660627)

[**Introducing “Golden Basket” 4**](#_Toc193660628)

[**Technical Details: RUG, LLM, Regression Model in Action 4**](#_Toc193660629)

[**RAG Architecture Overview (TinyLlama + FAISS): 5**](#_Toc193660630)

[**LLM – Large Language Model: 5**](#_Toc193660631)

[**Regression Model for Risk Scoring: 7**](#_Toc193660632)

[**Personalized Investment Basket Recommendation Engine 7**](#_Toc193660633)

[**GenAI-Powered Customer Interaction 7**](#_Toc193660634)

[**Conclusion and Future Roadmap 9**](#_Toc193660635)

**Golden Basket – An Automated Gen AI Powered Wealth Advisor**

**Summary**

In a rapidly evolving financial landscape, customer expectations from their banking partners have transformed significantly. They now seek **not just transactional services, but holistic financial experiences** that guide them toward financial well-being and wealth creation. The vision of "Golden Basket" is to offer a **GenAI-powered, highly personalized and customized wealth management platform** that enhances the banking experience and enables customers to make informed investment decisions.

Golden Basket acts as a **digital financial companion** that understands the customer's financial dreams and risk appetite, offering tailored recommendations in stocks, mutual funds, and investment baskets using AI-driven insights. The integration of advanced technologies such as **Risk Understanding Graphs (RUG)**, **Large Language Models (LLMs)**, and **regression-based risk profiling** ensures precision and personalization.

# **Target Audience:**

* Golden Basket is designed for a diverse audience of stakeholders in the financial ecosystem:
* Retail Banking Customers: Individuals who want to invest but lack time or expertise.
* HNIs (High Net Worth Individuals) & U-HNIs: Looking for intelligent portfolio diversification.
* Bank Wealth Managers: Need AI assistance in recommending personalized portfolios.
* Product Heads in Banks: Exploring next-gen digital banking innovations.
* Digital Transformation Leaders: Seeking GenAI-powered customer engagement models.
* Regulators & Policy Makers: Interested in how AI ensures customer-centric and responsible investing.

This ideation offers a roadmap to how AI can humanize digital wealth management and help customers to build their dreams, while being scalable, compliant, and transparent.

# **Problem Statement – Why Change is Needed?**

Before onboarding onto the investment portal of a Bank or NBFC, every customer undergoes a risk assessment, and investment recommendations are provided based on their risk appetite.

However, the current recommendation engine applies a standardized approach, offering uniform suggestions to customers with similar risk scores rather than catering to individual preferences.

To address this, there is a need for a hyper-personalized recommendation engine that delivers investment suggestions uniquely tailored to each customer's specific profile and needs.

**Key Challenges:**

**Generic Advice:** Most banks provide a one-size-fits-all product recommendation.

**Information Overload:** Customers struggle to analyse stock/fund data from various sources.

**Lack of Personalization:** Risk tolerance, goals, and life events are not adequately considered.

**Low Engagement:** Traditional platforms fail to spark emotional connection or confidence.

**Trust Deficit in AI:** Most customers don’t trust black-box AI recommendations.

Golden Basket addresses these problems by creating **tailored investment baskets with explainable AI-backed recommendations**, increasing both engagement and financial success.

# **Current Market Landscape**

Most wealth management platforms today follow:

* Rule-based static recommendations
* Generic Risk Profiling using few attributes
* Minimal use of AI/ML
* No GenAI narrative generation or customer conversation intelligence

Few institutions leverage AI in a truly explainable and personalized way. Robo-advisors are either too mechanical or too generic. There’s an urgent need for **emotionally intelligent, context-aware, personalized wealth tools**.

That’s where Golden Basket stands out — **not just as a recommendation engine, but a financial storytelling and advisory assistant**, leveraging **deep personalization through GenAI**.

# **Our Solution :**

## **Introducing “Golden Basket”**

Golden Basket is a **GenAI-Powered Personalized Wealth Management system** that enhances the banking experience and delivers customized investment solutions.

**Key Features:**

* Risk Score Prediction using Regression Modelling
* Personalized Investment Basket generation
* Natural Language Explanation using LLM
* Customer Need Analysis with RUG (Risk Understanding Graph)
* Investment Goal Mapping
* AI Storytelling – Explainable portfolio rationale
* Multi-Channel Delivery – Mobile, Web, Branch Advisor Assistance

**Customer Journey:**

1. Customer enters **goal, investment amount, time horizon**.
2. GenAI assistant **understands customer persona, risk profile, liquidity needs**.
3. System recommends **tailored asset baskets** (e.g., Tech Growth, Retirement Shield, Balanced Returns).
4. LLM **generates easy-to-understand investment rationale**.
5. Customer tracks performance, adjusts basket dynamically with AI guidance.

A screenshot of a chat

AI-generated content may be incorrect.

**Risk Score:**

*A risk score is a numerical representation of how likely a customer is to default on a loan, commit fraud, or behave in a high-risk financial manner. It helps financial institutions make data-driven decisions — whether it's approving a loan, offering insurance, or suggesting investments.*

In our use case , risk assessment is crucial for making informed decisions about loans, credit approvals, investments, and fraud detection. A risk score helps predict the likelihood of a customer making poor investment decisions.

By leveraging key customer attributes like age, geography, credit score, and mortgage details, we can build a comprehensive risk model that is both predictive and data-driven

| **Attribute** | **Why it Matters** |
| --- | --- |
|  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Age** | Younger customers may can take some risk when considering their time horizon; older ones need a stable income | |  |  | | |  |  | | --- | --- | | **Geography** | Certain areas may have high fraud rates, economic instability, or default trends .Also the value of immovable asset is proportional to the geographical area | |
| |  |  | | --- | --- | | **Credit Score** | Core measure of financial discipline and repayment behaviour. | | |  |  | | --- | --- | | **Mortgage Status** | Tied to liabilities; heavy mortgages mean more financial burden (higher risk). | |
| |  |  | | --- | --- | | **Employment Type** | Permanent employees in reputed firms are lower risk than freelancers or contractors. | | |  |  | | --- | --- | | **Income Level** | Higher income usually indicates better repayment capacity. | |
| |  |  | | --- | --- | | **Account Behaviour** | Transaction patterns (spending/saving habits) can indicate risk exposure. | | |  |  | | --- | --- | | **Previous Defaults** | Past behaviour is a strong predictor of future default probability. | |

Risk scoring systems are often built using machine learning models trained on historical customer data.These models find patterns and correlations between attributes and default probability.

Each attribute contributes to a **risk model**, often using **Machine Learning or Scoring Algorithms** (e.g., Logistic Regression, Random Forest, Gradient Boosting).

A **weighted score** is assigned to each factor. Here’s a **simplified formula** for a risk score:

Risk Score=(w1×Age Factor)+(w2×Geography Risk)+(w3×Credit Score)+(w4×Mortgage Factor) etc..

Where:

* **w1, w2, w3, w4** = Weightage based on historical data & business rules.
* **Age Factor** = Risk weight assigned to customer’s age.
* **Geography Risk** = Regional risk profile score.
* **Credit Score** = Standardized credit rating score.
* **Mortgage Factor** = Debt-to-income ratio based on mortgage obligations.

**Example**: A 28-year-old in a Tier-2 city with a low credit score and a heavy mortgage is statistically more likely to default and we cannot recommend high risk stocks/funds

We have a regression model to calculate the risk score, which is hyper tuned and trained periodically. We are supporting different personas and the model will give the risk score accordingly

**Once the risk score is received in the Recommendation System, it maps the score to a predefined asset allocation model**

**Model for Risk Score:**

Here the relationship between input features (e.g., customer financial data, credit score, investment history) and the risk score is highly non-linear, **RandomForestRegressor** is the choice we have selected. Unlike linear regression, it can capture intricate patterns and feature interactions. It handles high-dimensional datasets well and can identify which features are most influential.

* Random Forest provides feature importance scores, which help in understanding what factors contribute most to the risk score.
* This is useful for regulatory compliance, internal decision-making, or refining risk assessment models.
* Also, **Random Forest Regressor** can handle missing values better than many other models because it uses multiple trees and can make predictions even if some features are missing.

**Hyper Parameter Tuning:**

Hyperparameter tuning is crucial for improving the performance of a RandomForestRegressor, and GridSearchCV is one of the most systematic and widely used approaches for this. One major feature is it **Automates Model Evaluation**

* Instead of manually testing different parameter values, GridSearchCV automates the process.
* It trains multiple models using cross-validation and selects the combination that performs best.
* By selecting the best hyperparameters based on cross-validation scores, GridSearchCV prevents overfitting to a specific dataset.

**A diagram of a flowchart

AI-generated content may be incorrect.**

# 

# **Technical Details: RUG, LLM, Regression Model in Action**

## **RAG Architecture Overview (TinyLlama + FAISS):**

RAG stands for “Retrieval-Augmented Generation.” It’s a powerful architecture used to enhance the performance and accuracy of Large Language Models (LLMs) .

RAG = Retrieval + Generation

**Retrieval**: Before generating an answer, the system retrieves relevant information/documents from an external knowledge base, database, or document store (e.g., PDFs, websites, internal files).

**Augmentation**: This retrieved information is passed into the LLM as context.

**Generation**: The LLM then generates a response using both the user’s query and the retrieved information, leading to more accurate and contextually rich response

***User Query***

***↓***

***Embedding Model (e.g., all-MiniLM)***

***↓***

***Vector Search (FAISS over documents)***

***↓***

***Top-K Relevant Chunks***

***↓***

***Prompt = [Context + User Query]***

***↓***

***TinyLlama (Generates answer)***

**Stock Data**

To fetch stock ticker, we used - yfinance . It is a Python library that allows you to fetch real-time and historical stock market data from Yahoo Finance.

It is used for: Stock price analysis

* Fetching company financials (e.g., P/E ratio, market cap, earnings)
* Portfolio management
* Algorithmic trading

**Sentimental Analysis:**

We use - Yahoo Finance News Page

This provides:

* Latest News about the stock
* Market Trends impacting the company
* Earnings Reports & Analyst Ratings
* Mergers, Acquisitions, and Industry Insights

We uses LLM (Llama) to analyse sentiment.

* Counts how many news articles are Positive, Negative, or Neutral for a stock.
* This can be used to gauge investor sentiment based on recent news.

**Stock Analysis**

Stock ranking involves **assigning a score** to each stock based on multiple financial factors (e.g., price movements, volume, volatility, earnings). LGBMRegressor can be trained to **predict future stock returns** or an overall performance score, which is then used for ranking.

We have used LightGBM Regressor for stock analysis before recommendation. GBMRegressor (LightGBM Regressor) is a gradient boosting-based machine learning model optimized for speed and efficiency. It is part of the LightGBM library, which is designed for high-performance regression and classification tasks.

**Rule Engine**

A Rule Engine in stock recommendation is a logic-based system that applies predefined rules to recommend stocks based on various financial, technical, and sentiment factors. Unlike machine learning models, a rule engine explicitly encodes decision-making criteria and executes them dynamically to filter and rank stocks based on risk score and volatility based on customer profile.

For stock recommendations, the rules we used is based on

* **Fundamental Analysis** (e.g., P/E ratio, earnings growth)
* **Technical Analysis** (e.g., Moving Averages)
* **Sentiment Analysis** (e.g., News sentiment, social media trends)
* **Risk & Portfolio Diversification** (e.g., Avoid overexposure to a single sector)

Sample Rule:

* Diversify portfolio by selecting at least **3 different sectors**.
* Risk Matching Rule

Investor vs. Stock Volatility Risk Matching is the process of aligning an investor’s risk appetite with the volatility of a stock to ensure suitable investment choices. Investor Risk Score → Measures how much risk an investor is willing to take. Stock Volatility (Standard Deviation, Beta) → Measures how much the stock price fluctuates. Risk Matching → Ensures the investor is not exposed to unnecessary risk

**Rule Engine Output:**

Investor-Stock Scoring System assigns a score to each stock based on how well it matches an investor's profile using rules on:

**Investor Factors** like Net Worth, Liquidity, Risk Score, Region, Assets and Mortgage Debt

**Stock Factors like** Volatility, P/E Ratio, Market Cap, Sharpe Ratio, Sector Match

## **LLM – Large Language Model:**

We use LLM Model TinyLlama

This is used to

* Convert AI insights into human-readable language
* Provide **portfolio storytelling**
* Answer **"Why is this fund for me?"** in simple terms
* Explain **"What if" scenarios** in goal planning

*E.g., "We recommended this Tech Growth Basket because it aligns with your moderate risk appetite, and your time horizon of 5 years favors long-term capital appreciation."*

**Let’s say a customer asks:**

“What’s the stock/fund I need to include in my investment profile for a yield of 9% over the span of 10 years to achieve my financial goal?”

* In a **pure LLM model**, it might try to guess based on training data.
* In a **RAG model**, it:
  1. Retrieves recent and accurate information about that fund from a financial database.
  2. Feeds that into the LLM.
  3. LLM then generates a natural language response like:

**Sample Response:**

**A close-up of a document

AI-generated content may be incorrect.**

**The model recommends diversified sectorsA pie chart with numbers and a few percentages

AI-generated content may be incorrect.**

## **Personalized Investment Basket Recommendation Engine**

Each customer gets a **Golden Basket** which contains a blend of:

* Stocks
* Mutual Funds
* ETFs
* Gold/Crypto (optional based on user preference)

**Basket is generated based on:**

* Risk Score
* Time Horizon
* Sector Preference
* Liquidity Needs

## **GenAI-Powered Customer Interaction**

Golden Basket is more than an engine — it’s a **Digital Wealth Companion**.

**Features**:

Conversational AI Advisor (powered by LLM): "Hey Rahul, based on your current goals, would you like to rebalance your portfolio?"

Dream Goal Planner: "Help me save for my child’s education in 10 years."

Scenario Simulator: "What happens if I increase investment by ₹5,000/month?"

GenAI makes wealth management interactive, immersive, and human-like, which traditional platforms lack.

**ROI - Benefits :**

|  |  |
| --- | --- |
| Benefits to Customers: | Benefits to Bank: |
| Personalized Financial Empowerment | Deep customer engagement |
| Easy-to-understand investment logic | Nudge-based upsell and rebalancing |
| Confidence in AI suggestions | AI insights for advisors |
| Journey toward dream fulfilment | Regulatory-friendly explainable AI architecture |
| Increase wallet share and cross-sell |  |

**Business Value Differentiator**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Feature | **Traditional** **Platforms** | **Golden Basket** |
| Risk Profiling | Basic | Dynamic AI-based |
| Communication | Static | Conversational |
| Recommendation | Fixed list | Custom Basket |
| Explanation | Minimal | LLM-generated stories |
| Adaptability | Manual | Self-learning AI |

# **Challenges:**

**As with any advanced technological solution, implementing a highly personalized Gen-AI driven wealth management platform like Golden Basket comes with its own set of challenges.**

While the solution promises enhanced customer experience and intelligent portfolio recommendations, it also requires overcoming several technical complexities behind the scenes. From integrating large-scale financial datasets to ensuring real-time, context-aware responses via Retrieval-Augmented Generation (RAG) and Large Language Models (LLMs), each layer of the platform presents unique development, scalability, and compliance hurdles.  
 This section highlights the key technical challenges encountered in building and operationalizing the Golden Basket platform — challenges that not only test the robustness of AI infrastructure, but also its alignment with enterprise-grade banking standards, regulatory frameworks, and customer trust.

* Choosing between hosted vs self-hosted LLMs (e.g., OpenAI vs Mistral vs LLaMA). LLMs like GPT-4 have token limits (e.g., 8k/32k) — we can't pass unlimited RAG context into it. We must optimize prompt + retrieval.
* We need to design consistent prompts that include user intent, retrieved context, and formatting instructions (table/chart/text).
* Choosing Embedding Model - OpenAI, Cohere, HuggingFace models vary in quality, language support, and cost.

|  |
| --- |
| * FAISS works for small scale, we faced challenges when using for large scale data |

|  |
| --- |
| * Retrieval + LLM generation + formatting must happen in milliseconds — this is tricky when we use personnel Laptops for development. * Model Fusion - Combining regression risk scoring with RAG retrieval and LLM generation was bit challenging * Optimizing compute resources for regression scoring + retrieval + generation pipelines was a big challenge |

# **Conclusion and Future Roadmap**

Golden Basket is not just a product — it’s a **paradigm shift in digital wealth management**. By **combining AI’s precision with human-like communication**, it turns investments into stories and dreams into achievable goals by increasing customer engagement

**Future Roadmap:**

* Partner with stock exchanges for real-time data feeds
* Expand into insurance recommendation & tax advisory
* Build cross-sell opportunities via LLM-based nudges

**Golden Basket will redefine how customers perceive banks – not just as service providers, but as true financial partners.**