**KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG AI**

I. CẢI TIẾN DOLPHIN CHATBOT AI

1. Thực trạng vấn đề

* Chatbot hiện chỉ trả lời được một số câu hỏi chứng khoán cơ bản, thiếu linh hoạt với câu hỏi phức tạp.
* Khó phát triển sâu (phân tích kỹ thuật, cơ bản, tư vấn...) vì thiếu mô hình NLP chuyên biệt, chưa có kiến trúc phân tầng.
* Hiệu suất xử lý chưa cao, phản hồi chậm khi lượng người dùng lớn.
* Độ chính xác câu trả lời còn thấp, thiếu dẫn chứng, trích nguồn.
* Chưa áp dụng microservice, MLOps, hạn chế khả năng mở rộng.
* Chưa có quy trình tiền xử lý, kho dữ liệu vector, chưa quản lý lịch sử truy vấn và dữ liệu người dùng tập trung.
* Bảo mật & tuân thủ chưa hoàn thiện.

1. Giải pháp và phạm vi áp dụng

2.1. Giải pháp tổng quan

* Xây dựng kiến trúc microservice: tách biệt Data Pipeline, LLM Backend, Database, Frontend/Voice.
* Thiết lập MLOps: tự động huấn luyện, triển khai, giám sát mô hình.
* Tăng độ chính xác bằng cách cập nhật dữ liệu chứng khoán, bổ sung dẫn chứng, validator, feedback loop.
* Tối ưu hiệu suất: chia chatbot thành mô hình riêng cho từng chức năng, kết hợp caching, streaming, scale hạ tầng.
* Cải thiện bảo mật với Security & Compliance, mã hóa, kiểm soát quyền truy cập.

2.2. Chi tiết tính năng chatbot

* Truy vấn dữ liệu chứng khoán, chỉ số thị trường, tin tức.
* Hỗ trợ phân tích kỹ thuật (chỉ báo RSI, MACD...) và phân tích cơ bản (EPS, P/E, ROE...).
* Tóm tắt và phân tích tin tức thị trường, cung cấp sentiment.
* Lọc cổ phiếu theo tiêu chí, phân tích danh mục cá nhân, gợi ý tái cân bằng.
* Hỗ trợ thủ tục mở tài khoản, giao dịch, tư vấn kiến thức đầu tư, cảnh báo nội dung nhạy cảm.

2.3. Phạm vi áp dụng

* Nền tảng: Website, mobile, API (Zalo OA, FB Messenger...).
* Ngôn ngữ chính: Tiếng Việt (có thể mở rộng tiếng Anh).
* Đối tượng: Nhà đầu tư, tư vấn viên, CSKH.

2.4. Mục tiêu

* Thời gian phản hồi trung bình < 2 giây, độ chính xác > 80%.
* Giảm 50% khối lượng công việc CSKH, tăng chỉ số hài lòng khách hàng.

1. Kế hoạch phát triển

3.1. Lộ trình (5 tháng)

* Giai đoạn 1 (Tháng 1): Khởi tạo và thiết kế. Phân tích hệ thống cũ, chọn công nghệ, triển khai khung MLOps.
* Giai đoạn 2 (Tháng 2 - 3): Xây dựng data pipeline, tích hợp caching, Websocket, phát triển các module cốt lõi (truy vấn, phân tích kỹ thuật/cơ bản, tóm tắt tin).
* Giai đoạn 3 (Tháng 4): Hoàn thiện mô hình nâng cao (Legal Supporter, Investment Advisor), tối ưu hiệu suất & bảo mật.
* Giai đoạn 4 (Tháng 5): UAT, A/B testing với chatbot cũ, sửa lỗi, triển khai chính thức, thu thập phản hồi người dùng.

3.2. Kiểm thử và quản lý chất lượng

* Kiểm thử chức năng, hiệu suất (stress test), bảo mật (kiểm tra lỗ hổng).
* A/B Testing với chatbot cũ để so sánh cải tiến.

3.3. Quản lý rủi ro

* Dữ liệu thiếu, lỗi thời → Xây dựng pipeline tin cậy.
* Tăng đột biến người dùng → Kubernetes autoscale.
* Thông tin sai lệch → Thêm validator, feedback loop.
* Rò rỉ dữ liệu → Mã hóa thông tin nhạy cảm, kiểm soát bảo mật.

3.4. Phương hướng cải tiến (Sau 5 tháng)

* Hỗ trợ đa ngôn ngữ, tích hợp Voice Assistant, gợi ý câu hỏi liên quan, hành động thông minh (watchlist, cảnh báo).
* Tự động retrain, giám sát drift, kết nối đa kênh (Zalo OA, Messenger, Email…).

**KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG AI NEWS**

1. Thực trạng vấn đề

* Công ty nhận lượng tin tức khổng lồ mỗi ngày, gây khó khăn khi phân tích thủ công.
* Chưa có hệ thống phân loại chủ đề, phân tích sentiment, trích xuất mã cổ phiếu, dẫn đến khó nắm bắt xu hướng thị trường.
* Chưa có công cụ hỗ trợ ra quyết định, thiếu dashboard hay mô hình ML phân tích chuyên sâu.
* Người dùng muốn “Tin tức thành Podcast” nhưng chưa có tính năng TTS.

1. Giải pháp và phạm vi áp dụng

2.1. Giải pháp tổng quan

* Xây dựng pipeline NLP: Thu thập tin, tiền xử lý, phân tích sentiment, phân loại ngành, trích xuất mã cổ phiếu, tóm tắt, tính điểm cảm xúc...
* Áp dụng mô hình ngôn ngữ (BERT, GPT…) cho Sentiment, Topic Segmentation, NER, Summarization, TTS.
* Tích hợp dữ liệu mạng xã hội để đánh giá quan điểm cộng đồng.
* Xây dựng MLOps giúp tự động cập nhật, retrain mô hình, giám sát hiệu năng.

2.2. Chi tiết tính năng

* Xác định sentiment (tích cực/tiêu cực/trung tính), phân loại theo ngành, trích xuất mã cổ phiếu.
* Tóm tắt tin tức, tính điểm sentiment cho cổ phiếu theo mốc thời gian.
* Phân tích dữ liệu mạng xã hội (social sentiment).

2.3. Phạm vi áp dụng

* Nguồn dữ liệu: API tin tức, crawler, mạng xã hội.
* Đối tượng: Bộ phận phân tích, tư vấn đầu tư, marketing, nhà đầu tư.
* Kênh: Web/App, tích hợp API, dashboard real-time.

2.4. Mục tiêu

* Xử lý tin tức trong vòng 10 giây, sentiment F1-score 80-85%, phân loại ngành ≥85% chính xác, tóm tắt ROUGE/BLEU ≥0.6-0.7.
* Giảm 50% khối lượng phân tích thủ công, triển khai đầy đủ trong 2 tháng.

1. Kế hoạch phát triển 3.1. Lộ trình (2 tháng)

* Giai đoạn 1: Thiết lập hạ tầng, data pipeline, tiền xử lý (loại trùng lặp, làm sạch).
* Giai đoạn 2: Huấn luyện mô hình sentiment, phân loại chủ đề, NER, tóm tắt tin tức, tính điểm cảm xúc cổ phiếu, kiểm thử dữ liệu thực tế.
* Giai đoạn 3: Tích hợp, tối ưu, xây dashboard phân tích, tăng tốc pipeline, triển khai cảnh báo sentiment.
* Giai đoạn 4: Hoàn thiện, phát triển TTS, kiểm tra bảo mật, tinh chỉnh mô hình, thử nghiệm người dùng, go-live chính thức.

3.2. Kiểm thử và quản lý chất lượng

* Kiểm thử chức năng pipeline, hiệu suất (throughput/latency), độ chính xác sentiment, topic, tóm tắt.
* Bảo mật dữ liệu, ẩn danh, A/B testing so với hệ thống cũ.

3.3. Quản lý rủi ro

* Dữ liệu kém chất lượng → Cần cơ chế làm sạch & kiểm soát.
* Tin tức quá tải → Xây batch/streaming, scale hạ tầng.
* Sai sót mô hình → Giám sát và retrain khi cần.
* Thiếu nhân sự chuyên sâu → Phân chia task, đảm bảo đủ nguồn lực.

3.4. Phương hướng cải tiến

* Tích hợp thêm nhiều nguồn tin, tính năng gợi ý tin quan trọng, phát triển mảng podcast, phân tích quan hệ sự kiện, chatbot AI tương tác thời gian thực, cảnh báo sentiment đột biến.

**KẾ HOẠCH PHÁT TRIỂN MÔ HÌNH PHÂN TÍCH KHÁCH HÀNG (Phân loại khách hàng, dự đoán churn, tìm kiếm khách hàng tiềm năng – Thời gian: 3 tháng)**

I. CẢI TIẾN HỆ THỐNG PHÂN TÍCH KHÁCH HÀNG

1. Thực trạng vấn đề

* Thiếu phân khúc khách hàng rõ ràng, khó triển khai marketing nhắm đúng nhóm.
* Tỷ lệ rời bỏ dịch vụ (churn) cao, chưa có công cụ dự đoán.
* Nguồn khách hàng từ công ty mẹ lớn nhưng chưa có mô hình lead scoring.
* Dữ liệu phân tán, chưa chuẩn hóa, gây khó khăn cho phân tích tổng thể.

II. GIẢI PHÁP VÀ PHẠM VI ÁP DỤNG

2.1. Giải pháp tổng quan

* Mô hình phân loại khách hàng (Customer Segmentation) bằng RFM, K-Means…
* Mô hình dự đoán churn (Logistic Regression, XGBoost…).
* Mô hình lead scoring: dùng hồ sơ tài chính, hành vi giao dịch để xếp hạng tiềm năng.
* Kết nối với Feature Store để dùng chung dữ liệu, tránh trùng lặp.

2.2. Chi tiết tính năng

* Segmentation: Chia nhóm VIP, tiềm năng, ngủ đông… + dashboard đặc điểm từng nhóm.
* Churn Prediction: Dự báo khách sắp rời bỏ, cảnh báo đội CSKH/marketing.
* Lead Scoring: Ưu tiên tiếp cận khách hàng tiềm năng từ danh sách công ty mẹ.

2.3. Phạm vi áp dụng

* Dữ liệu từ CRM, lịch sử giao dịch, hành vi web/app, khách hàng công ty mẹ.
* Bộ phận marketing, CSKH, sales.
* Mô hình batch inference, có thể mở rộng real-time.

2.4. Mục tiêu

* Giảm 20% churn, tăng 15% chuyển đổi khách hàng tiềm năng, phân khúc rõ ràng.
* Hoàn thành trong 3 tháng (tích hợp vào quy trình kinh doanh).

III. KẾ HOẠCH PHÁT TRIỂN

3.1. Lộ trình (3 tháng)

* Giai đoạn 1: Thu thập, chuẩn hóa dữ liệu (CRM, giao dịch, logs...), xác định KPI churn/lead scoring.
* Giai đoạn 2: Huấn luyện mô hình segmentation (K-Means, RFM), churn (LR, XGBoost), lead scoring, chọn mô hình tốt nhất.
* Giai đoạn 3: Tích hợp mô hình vào quy trình CSKH/marketing, xây dashboard, kiểm thử thực tế, lấy phản hồi.
* Giai đoạn 4: Tối ưu, triển khai chính thức, thiết lập pipeline MLOps, A/B test đánh giá kết quả.

3.2. Kiểm thử & Quản lý chất lượng

* Segmentation: Kiểm tra trực quan, tính ổn định cụm.
* Churn: Dùng ROC-AUC, F1-score.
* Lead Scoring: so sánh tỷ lệ chuyển đổi.
* Triển khai thử trên một phần khách hàng, kiểm tra bảo mật dữ liệu.

3.3. Quản lý rủi ro

* Chưa có nhãn churn rõ ràng → Cần thống nhất định nghĩa.
* Dữ liệu công ty mẹ không đầy đủ → Xử lý thiếu, chuẩn hóa.
* Tính hiệu quả mô hình không như mong đợi → Tối ưu hyperparameter, cập nhật liên tục.

3.4. Phương hướng cải tiến

* Mở rộng sang mô hình upsell/cross-sell, đề xuất sản phẩm, cá nhân hóa marketing.
* Online learning cho churn/lead scoring real-time.
* Thêm phân tích xu hướng theo thời gian, deep learning để nâng cao dự báo.

**KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG PHÂN TÍCH VOC (VOICE OF CUSTOMER) & TICKET (Phân loại và đánh nhãn sentiment – Thời gian: 2 tháng)**

I. CẢI TIẾN HỆ THỐNG PHÂN TÍCH VOC & TICKET

1. Thực trạng vấn đề

* Số lượng VOC & ticket lớn, cần phân loại thủ công tốn thời gian.
* Chưa có đánh giá sentiment tự động (tiêu cực, trung tính, tích cực).
* Dữ liệu thô, chưa phân loại theo chủ đề (phí, kỹ thuật, khiếu nại…).
* Định hướng tích hợp LLM API để phân loại, phân tích nội dung.

II. GIẢI PHÁP VÀ PHẠM VI ÁP DỤNG

2.1. Giải pháp tổng quan

* Thu thập & tiền xử lý VOC/ticket: trích nội dung từ email, chat, ghi âm (STT).
* Phân loại chủ đề bằng LLM API hoặc mô hình fine-tuned.
* Phân tích sentiment tự động, sắp xếp mức ưu tiên xử lý.
* Sử dụng prompt engineering, nhận kết quả dạng JSON, có feedback loop nâng cao độ chính xác.

2.2. Chi tiết tính năng

* Tự động gán “Loại vấn đề”, “Độ ưu tiên”.
* Phân tích cảm xúc, cảnh báo khi sentiment quá tiêu cực.
* Dashboard theo dõi tỷ lệ VOC tích cực/tiêu cực, các chủ đề phổ biến.

2.3. Phạm vi áp dụng

* Nguồn VOC: email, chat, form phản hồi, call (chuyển text).
* Ticket: Tích hợp với Zendesk/ServiceNow/Jira hoặc hệ thống nội bộ.
* Bộ phận sử dụng: CSKH, quản lý chất lượng.

2.4. Mục tiêu

* Tự động phân loại ≥80% ticket/VOC, sentiment F1-score ≥85%.
* Giảm 50% thời gian sàng lọc thủ công, hoàn thành MVP trong 2 tháng.

III. KẾ HOẠCH PHÁT TRIỂN

3.1. Lộ trình (2 tháng)

* Giai đoạn 1: Xác định các chủ đề, khung sentiment, thu thập & chuẩn hóa dữ liệu VOC/ticket.
* Giai đoạn 2: Thiết kế prompt/API call cho LLM, phát triển prototype, test trên dữ liệu thực tế.
* Giai đoạn 3: Tích hợp và tối ưu, xây dashboard, thiết lập feedback loop để CSKH điều chỉnh nhãn.
* Giai đoạn 4: Kiểm thử cuối, đưa vào hệ thống ticketing chính, đào tạo đội CSKH, go-live.

3.2. Kiểm thử & Quản lý chất lượng

* Đo F1-score, Precision, Recall giữa nhãn LLM và nhãn thủ công.
* Kiểm thử hiệu suất (thời gian xử lý, chịu tải), bảo mật (ẩn danh dữ liệu).
* Thử nghiệm trên một vài phòng ban trước khi mở rộng toàn công ty.

3.3. Quản lý rủi ro

* Chi phí gọi LLM API cao nếu volume lớn → Cần batch processing, cache.
* Chủ đề hiếm → Thu thập thêm ví dụ để mô hình nhận diện.
* Bảo mật dữ liệu khách hàng → Mã hóa, ẩn danh trước khi gửi API.
* API downtime → Chuẩn bị queue, cơ chế fallback.

3.4. Phương hướng cải tiến

* Auto-reply cho câu hỏi/issue đơn giản.
* Hỗ trợ đa ngôn ngữ, phân tích voice sentiment trực tiếp từ ghi âm.
* Tùy chọn dùng LLM on-prem với dữ liệu nhạy cảm.

**KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG FEATURE STORE CHO CÁC BÀI TOÁN BUSINESS (Thời gian triển khai: 3 tháng)**

I. CẢI TIẾN HỆ THỐNG QUẢN LÝ FEATURE

1. Thực trạng vấn đề

* Dữ liệu phân tán: CRM, giao dịch, công ty mẹ… lưu ở nhiều nơi.
* Mỗi nhóm phân tích tự làm sạch, chuẩn hóa riêng, gây lặp lại, thiếu “single source of truth”.
* Khó cập nhật feature khi mở rộng mô hình.
* Triển khai mô hình vào sản xuất chậm vì ETL thủ công, dữ liệu không đồng nhất.

II. GIẢI PHÁP VÀ PHẠM VI ÁP DỤNG

2.1. Giải pháp tổng quan

* Xây dựng Feature Store quản lý tập trung toàn bộ feature (batch + near real-time).
* Quản lý version, cho phép time-travel, cung cấp API/SDK để truy cập nhất quán.
* Kết hợp MLOps: tự động hóa pipeline, giám sát chất lượng dữ liệu, drift, cảnh báo khi có lỗi.

2.2. Chi tiết tính năng

* Feature Catalog: Liệt kê đầy đủ feature (định nghĩa, logic tính toán, ngày cập nhật).
* Versioning & Time Travel: Phục hồi version tính năng đúng với thời điểm training.
* Serving layer: batch (định kỳ) và online (gần real-time).
* Metadata & Access Control: Quản lý mô tả, cấp quyền, bảo mật tính năng nhạy cảm.

2.3. Phạm vi áp dụng

* Dữ liệu: CRM, lịch sử giao dịch, dữ liệu công ty mẹ, đối tác…
* Ứng dụng: churn prediction, segmentation, lead scoring, khuyến nghị sản phẩm…

2.4. Mục tiêu

* Giảm 50% thời gian chuẩn bị dữ liệu cho các dự án ML, đảm bảo tính nhất quán.
* Hỗ trợ batch/online, triển khai MVP trong 3 tháng.

III. KẾ HOẠCH PHÁT TRIỂN

3.1. Lộ trình (3 tháng)

* Giai đoạn 1: Phân tích yêu cầu, thiết kế kiến trúc, chọn công nghệ (Feast, Databricks...), xây dựng pipeline ETL batch ban đầu, PoC Feature Store cho một số tính năng.
* Giai đoạn 2: Mở rộng ingestion (dữ liệu công ty mẹ, logs...), tích hợp versioning/time travel, kết nối MLOps, CI/CD, thử nghiệm với 1-2 mô hình thực tế (churn/lead scoring).
* Giai đoạn 3: Kiểm thử chất lượng feature, giám sát data quality, bảo mật, tinh chỉnh pipeline, đào tạo đội ngũ Data Science, go-live MVP và đo hiệu quả.

3.2. Kiểm thử & Quản lý chất lượng

* Kiểm thử chính xác: So sánh kết quả feature mới với pipeline cũ.
* Kiểm thử hiệu năng: Thời gian load, khả năng mở rộng.
* Kiểm thử bảo mật: Quy trình mã hóa, phân quyền.
* Pilot deployment: Tích hợp cho một số mô hình ML, theo dõi ROI.

3.3. Quản lý rủi ro

* Nhiều bên liên quan, dễ chậm trễ → Cần điều phối tốt.
* Sai sót dữ liệu → Validation, alert khi có bất thường.
* Hạ tầng phức tạp → DevOps có kinh nghiệm Feature Store.
* Chi phí vận hành → Kiểm soát nếu dùng cloud hay phần mềm trả phí.

3.4. Phương hướng cải tiến

* Tích hợp thêm nguồn dữ liệu web/app, social.
* Xây dựng pipeline streaming real-time, auto feature engineering gợi ý tính năng mới.
* Tích hợp chặt với MLOps để retrain, giám sát drift tự động.

**KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG GỢI Ý (3 THÁNG)**

I. CẢI TIẾN HỆ THỐNG GỢI Ý

1. Thực trạng vấn đề

* Hệ thống hiện chưa có khả năng đề xuất sản phẩm/tính năng phù hợp cho từng đối tượng người dùng (nhà đầu tư cá nhân, newbie, khách VIP...).
* Dữ liệu về hành vi (usage logs, clickstreams) và lịch sử giao dịch chưa được khai thác một cách đồng bộ để phục vụ gợi ý.
* Thiếu cơ chế linh hoạt để đề xuất:
  + Sản phẩm đầu tư (options, futures, margin...).
  + Gói tối ưu dòng tiền (bonds, short deposit...).
  + Tin tức phù hợp sở thích.
  + Danh mục cổ phiếu dựa trên tiêu chí lợi nhuận/rủi ro.
* Các chiến dịch marketing hiện chủ yếu vẫn được thực hiện theo nhóm lớn, chưa cá nhân hóa.
* Môi trường giao dịch chưa có tính năng “cá nhân hóa” (tự động gợi ý dashboard, công cụ, tính năng phù hợp).

1. Giải pháp và phạm vi áp dụng

2.1. Giải pháp tổng quan

* Áp dụng các phương pháp gợi ý: Collaborative Filtering (User-User, Item-Item), Content-based Filtering, Hybrid (Matrix Factorization + Deep Learning).
* Sử dụng Reinforcement Learning (RL) để điều chỉnh giao diện/trải nghiệm người dùng (UI) dựa trên mức độ tương tác và phản hồi.
* Tích hợp dữ liệu toàn diện (trade history, demographics, hành vi duyệt tin, clickstreams) vào kho dữ liệu chung để phân tích và huấn luyện mô hình.
* Phát triển pipeline MLOps giúp retrain mô hình định kỳ hoặc khi có dữ liệu mới.

2.2. Chi tiết tính năng gợi ý chính

1. Onboarding Newbie
   * Mục tiêu: Hướng dẫn nhà đầu tư mới làm quen với nền tảng, gợi ý các bước nhập môn, video hướng dẫn, tài liệu học cơ bản.
   * Dữ liệu cần: Thông tin đăng ký tài khoản, nhân khẩu học, hành vi truy cập ban đầu.
   * Phương pháp:
     + Content-based filtering để đề xuất tài liệu/hướng dẫn phù hợp.
     + Xây hành trình onboarding từng bước (step-by-step) dựa trên mức độ hoàn thành trước đó.
2. Service Recommendation
   * Other Investment Securities: Gợi ý các sản phẩm liên quan (options, futures, margin accounts), báo cáo phân tích, khóa học.
   * Cash Optimization (Bond, Short Deposit): Đề xuất gói tối ưu dòng tiền, sản phẩm trái phiếu, hoặc gửi ngắn hạn tùy theo số dư tài khoản, khẩu vị rủi ro.
   * News Recommendation: Đưa tin tức, bài phân tích phù hợp với danh mục và sở thích người dùng.
   * Dữ liệu cần: Lịch sử giao dịch, hồ sơ tài chính, mức độ quan tâm trước đó (browsing patterns), thông tin meta sản phẩm.
   * Phương pháp:
     + Collaborative Filtering (user-user, item-item) dựa trên hành vi, lượt xem và giao dịch của những người dùng tương tự.
     + Content-based filtering với embeddings (mô tả sản phẩm, chủ đề tin tức...).
3. Stock Recommendation
   * Đề xuất mã cổ phiếu tương tự với danh mục hiện tại (Similar Customer/Portfolio).
   * Gợi ý những cổ phiếu nổi bật (Outstanding Item) dựa trên xu hướng thị trường, phân tích thống kê.
   * Xây dựng “Portfolio Optimization” theo tiêu chí rủi ro, lợi nhuận, ngành, thanh khoản...
   * Dữ liệu cần: Danh mục cổ phiếu hiện có, lịch sử giao dịch, thông tin thị trường, hồ sơ rủi ro nhà đầu tư.
   * Phương pháp: Hybrid recommendation (kết hợp CF + phân tích kĩ thuật/cơ bản), các mô hình portfolio optimization (Markowitz, deep RL...).

2.3. Phạm vi áp dụng

* Nguồn dữ liệu:
  + Lịch sử giao dịch, hồ sơ người dùng (CRM, demographics).
  + Hành vi duyệt web/app (clickstream, thời gian đọc bài, độ tương tác).
  + Dữ liệu thị trường (bảng giá, tin tức, phân tích).
* Đối tượng sử dụng: nhà đầu tư cá nhân (newbie, trung cấp, cao cấp), các chuyên viên tư vấn, bộ phận marketing.
* Các kênh tích hợp: website giao dịch, mobile app, chatbot, dashboard quản trị.

2.4. Mục tiêu

* Tăng 20-30% tỷ lệ tương tác (click-through rate) của các gợi ý so với trước.
* Giảm tỷ lệ thoát (churn) của người dùng mới trong 3 tháng đầu.
* Thúc đẩy tỉ lệ sử dụng đa dạng dịch vụ (options, bonds...) ít nhất +15%.
* Tạo trải nghiệm cá nhân hóa và giữ chân khách hàng.
* Thời gian triển khai ban đầu: 3 tháng để có MVP đưa vào thử nghiệm.

**III. KẾ HOẠCH PHÁT TRIỂN**

3.1. Lộ trình triển khai (3 tháng)

Giai đoạn 1: Phân tích & Chuẩn bị dữ liệu (Tháng 1)

* Xác định danh sách tính năng gợi ý chi tiết (onboarding newbie, service recommendation, stock recommendation).
* Tập hợp và làm sạch dữ liệu: lịch sử giao dịch, profile khách hàng, clickstreams.
* Thiết lập kiến trúc hạ tầng (data lake/warehouse, pipelines, công cụ phân tích).
* Phát triển bản mẫu mô hình CF cơ bản hoặc content-based (PoC).

Giai đoạn 2: Xây dựng & Huấn luyện mô hình (Tháng 2)

* Hoàn thiện mô hình recommendation (CF, content-based, hybrid).
* Tích hợp thêm Reinforcement Learning (hoặc bandit) cho UI cá nhân hóa.
* Triển khai pipeline MLOps: tự động huấn luyện/retrain, theo dõi hiệu suất mô hình.
* Bắt đầu thử nghiệm nội bộ (internal UAT) với tập người dùng hạn chế.

Giai đoạn 3: Tích hợp, Kiểm thử & Triển khai (Tháng 3)

* Tích hợp hệ thống gợi ý vào website/mobile app, hiển thị gợi ý cho người dùng thật.
* Xây dựng dashboard cho đội marketing, tư vấn để theo dõi chất lượng gợi ý, mức độ tương tác.
* Thực hiện A/B Testing, so sánh mô hình mới với nhóm đối chứng (không gợi ý cá nhân hóa).
* Tối ưu thông số, sửa lỗi, điều chỉnh tính năng, go-live MVP rộng rãi.
* Thu thập phản hồi người dùng để tiếp tục cải tiến.

3.2. Kiểm thử và quản lý chất lượng

* Kiểm thử chức năng: Đảm bảo mô hình gợi ý chạy đúng, hiển thị phù hợp trong giao diện.
* Kiểm thử hiệu suất: Đo thời gian phản hồi gợi ý (latency), khả năng xử lý khi tăng đột biến người dùng.
* Kiểm thử độ chính xác: Sử dụng các chỉ số CTR (Click-Through Rate), Precision@K, Recall@K, nDCG...
* A/B Testing: So sánh trải nghiệm người dùng trước/sau khi bật gợi ý cá nhân hóa.
* Kiểm thử bảo mật: Đảm bảo ẩn danh dữ liệu người dùng, tuân thủ chính sách an toàn.

3.3. Quản lý rủi ro

* Dữ liệu không đồng nhất hoặc thiếu → Xây dựng quy trình làm sạch, chuẩn hóa, bù đắp thiếu (imputation).
* Hiệu quả mô hình không cao → Thử nghiệm nhiều thuật toán, thường xuyên kiểm tra & tối ưu hyperparameter.
* Tăng đột biến lượng truy cập → Thiết lập autoscale hạ tầng (Kubernetes, cloud).
* Bảo mật dữ liệu cá nhân (PII) → Áp dụng mã hóa, phân quyền, ẩn danh trước khi đưa vào mô hình.
* Phản ứng người dùng trái chiều với gợi ý → Cho phép tùy chọn “tắt gợi ý” hoặc tùy biến mức độ cá nhân hóa.

3.4. Phương hướng cải tiến (sau 3 tháng)

* Bổ sung tính năng Marketing Campaign Targeting:
  + Cá nhân hóa chiến dịch marketing dựa trên hành vi, phân khúc người dùng.
  + Theo dõi chuyển đổi cụ thể cho từng chiến dịch, tích hợp với hệ thống CRM/Email/SMS.
* Nâng cấp Personalization of Trading Platform:
  + Tự động sắp xếp dashboard, khuyến nghị công cụ phù hợp phong cách trading (kỹ thuật, cơ bản...).
  + Reinforcement Learning để cải thiện giao diện liên tục dựa trên tương tác thời gian thực.
* Mở rộng mô hình gợi ý cho các kênh khác (chatbot, voice assistant).
* Áp dụng Deep Learning (Transformer-based, seq2seq) cho việc gợi ý tin tức và sản phẩm phức tạp.
* Triển khai multi-armed bandit hoặc RL nâng cao để cá nhân hóa triệt để giao diện, ưu đãi, thông báo.
* Bổ sung mô hình Social Recommendation: khai thác dữ liệu mạng xã hội (nếu khả thi) để đánh giá xu hướng, cảm xúc công chúng.

**KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG GỢI Ý (4 THÁNG)**

I. CẢI TIẾN HỆ THỐNG GỢI Ý

1. Thực trạng vấn đề

* Nền tảng chưa có khả năng đề xuất sản phẩm, dịch vụ, nội dung phù hợp cho từng loại người dùng (newbie, nhà đầu tư dày dạn, khách VIP...).
* Dữ liệu về hành vi (usage logs, clickstreams, trade history) chưa được tích hợp đầy đủ để tận dụng cho bài toán gợi ý.
* Chưa có tính năng cá nhân hóa giao diện (UI), tư vấn sản phẩm đầu tư, hoặc tối ưu danh mục trên quy mô lớn.
* Chiến dịch marketing chủ yếu sử dụng phân khúc rộng, chưa “target” chính xác tới từng khách hàng.
* Cần bổ sung các hướng cải tiến trong tương lai (Marketing Campaign Targeting, Personalization of Trading Platform) sau khi hệ thống nền tảng hoạt động ổn định.

1. Giải pháp và phạm vi áp dụng

2.1. Giải pháp tổng quan

* Áp dụng các thuật toán recommendation: Collaborative Filtering (User-User, Item-Item), Content-based Filtering, Hybrid (kết hợp Matrix Factorization + Deep Learning).
* Sử dụng Reinforcement Learning (RL) hoặc multi-armed bandit để điều chỉnh giao diện dựa trên mức độ tương tác.
* Tích hợp dữ liệu từ nhiều nguồn (trade history, demographics, logs, tin tức…) vào kho dữ liệu thống nhất, phục vụ huấn luyện mô hình.
* Phát triển pipeline MLOps nhằm hỗ trợ quá trình huấn luyện, triển khai, giám sát mô hình một cách tự động.

2.2. Chi tiết tính năng gợi ý chính

1. Onboarding Newbie
   * Gợi ý lộ trình “nhập môn” giúp nhà đầu tư mới hiểu nền tảng, các tính năng cơ bản, tài liệu học…
   * Dữ liệu cần: Thông tin đăng ký, nhân khẩu học, hành vi truy cập ban đầu.
   * Phương pháp: Content-based filtering cho tài liệu, step-by-step tutorials theo hành vi tiến triển.
2. Service Recommendation
   * Other Investment Securities: Đề xuất sản phẩm liên quan (options, futures, margin), báo cáo phân tích, khóa học.
   * Cash Optimization: Gợi ý gói trái phiếu, gửi ngắn hạn… theo số dư tài khoản, khẩu vị rủi ro.
   * News Recommendation: Tìm kiếm, lựa chọn tin tức phù hợp danh mục/sở thích người dùng.
   * Dữ liệu cần: Giao dịch quá khứ, hồ sơ tài chính, hành vi duyệt nội dung, thông tin meta sản phẩm.
   * Phương pháp:
     + Collaborative Filtering (user-user, item-item) dựa trên hành vi.
     + Content-based embeddings để mô tả sản phẩm, tin tức.
     + Hybrid systems (kết hợp CF + deep learning).
3. Stock Recommendation
   * Similar Customer/Portfolio: Tìm mã cổ phiếu tương tự dựa trên danh mục hiện có, phân tích “người dùng tương tự”.
   * Outstanding Item: Gợi ý cổ phiếu nổi bật, xu hướng thị trường.
   * Portfolio Optimization: Đưa ra danh mục tối ưu theo tiêu chí rủi ro/lợi nhuận/ngành.
   * Dữ liệu cần: Danh mục cổ phiếu, lịch sử giao dịch, chỉ số thị trường, hồ sơ rủi ro.
   * Phương pháp: Hybrid recommendation + kỹ thuật portfolio optimization (Markowitz, deep RL…).

2.3. Phạm vi áp dụng

* Nguồn dữ liệu:
  + Giao dịch, CRM, lịch sử tương tác web/app, bảng giá, tin tức, logs sự kiện.
* Đối tượng sử dụng: Người dùng newbie, trung cấp, cao cấp; đội marketing, tư vấn.
* Kênh triển khai: Website, mobile app, chatbot, dashboard quản trị.

2.4. Mục tiêu

* Nâng cao tỷ lệ tương tác với các đề xuất (click-through rate) thêm 20-30%.
* Giảm churn của nhóm người dùng mới.
* Tăng ít nhất 15% việc sử dụng dịch vụ đầu tư đa dạng (options, bonds…).
* Cải thiện trải nghiệm cá nhân hóa, giữ chân khách hàng.
* Thời gian thực hiện: 4 tháng để đưa MVP vào thử nghiệm thực tế.

**III. KẾ HOẠCH PHÁT TRIỂN**

3.1. Lộ trình triển khai (4 tháng)

Giai đoạn 1 (Tháng 1): Phân tích & Chuẩn bị dữ liệu

* Thu thập yêu cầu chi tiết cho từng tính năng gợi ý (onboarding newbie, service recommendation, stock recommendation).
* Tổng hợp dữ liệu: lịch sử giao dịch, hồ sơ khách hàng, logs hành vi (clickstream), danh mục…
* Thiết lập nền tảng hạ tầng: data lake/warehouse, pipeline ETL cơ bản, lựa chọn công nghệ mô hình (framework, thư viện).
* Phát triển PoC (proof of concept) mô hình CF hoặc content-based cơ bản để đánh giá dữ liệu, tiềm năng.

Giai đoạn 2 (Tháng 2): Xây dựng & Huấn luyện mô hình

* Hoàn thiện mô hình recommendation: kết hợp CF, content-based, hybrid, thêm mô hình RL nếu cần.
* Thiết kế tính năng “gợi ý onboarding newbie” (các bước hướng dẫn, video, tài liệu).
* Tích hợp MLOps pipeline: tự động hóa huấn luyện, theo dõi hiệu suất (logging, metrics).
* Chạy thử nghiệm mô hình trên dữ liệu nội bộ, hiệu chỉnh tham số (hyperparameter tuning).

Giai đoạn 3 (Tháng 3): Tích hợp, Kiểm thử & Tối ưu

* Tích hợp hệ thống gợi ý với frontend website/app: tạo widget đề xuất dịch vụ, sản phẩm, cổ phiếu.
* Hoàn thiện UI/UX cho người dùng, bố trí vị trí hiển thị gợi ý (homepage, trang phân tích, trang cá nhân…).
* Kiểm thử A/B: so sánh với nhóm đối chứng để đo CTR, conversion, churn.
* Tối ưu hiệu suất hạ tầng: caching, batch vs. real-time inference, autoscaling (nếu cần).
* Xây dựng dashboard giám sát chất lượng gợi ý, theo dõi độ hài lòng người dùng.

Giai đoạn 4 (Tháng 4): Mở rộng & Go-live MVP

* Chính thức triển khai MVP cho nhóm người dùng lớn hơn, thu thập phản hồi.
* Bổ sung Service Recommendation chi tiết (options, futures, bonds...) dựa trên dữ liệu thực tế.
* Kết nối pipeline với dữ liệu thị trường thời gian thực (nếu khả thi) để gợi ý cổ phiếu nổi bật.
* Đào tạo đội ngũ marketing/CSKH, hướng dẫn sử dụng dashboard, khai thác báo cáo.
* Tổng kết hiệu suất, đánh giá ROI, lên kế hoạch cải tiến giai đoạn tiếp theo.

3.2. Kiểm thử và quản lý chất lượng

* Kiểm thử chức năng: Đảm bảo mô hình xuất gợi ý chính xác, UI hiển thị đúng.
* Kiểm thử hiệu suất: Đo latencies, throughput, sẵn sàng scale khi user tăng đột biến.
* Kiểm thử độ chính xác mô hình: Precision@K, Recall@K, nDCG, CTR…
* A/B Testing: Đo tỉ lệ chuyển đổi, mức độ tương tác với gợi ý.
* Kiểm thử bảo mật: Kiểm tra ẩn danh dữ liệu, tuân thủ chính sách an toàn.

3.3. Quản lý rủi ro

* Dữ liệu thiếu/chưa đồng nhất → Cần làm sạch, thống nhất, xử lý missing data.
* Hiệu quả mô hình thấp → Thử nghiệm nhiều thuật toán, cải thiện tính năng, thường xuyên theo dõi & tinh chỉnh.
* Tăng đột biến user → Xây dựng cơ chế load balancing, autoscaling (Kubernetes, cloud).
* Rủi ro bảo mật → Mã hóa dữ liệu nhạy cảm, phân quyền truy cập theo chuẩn.
* Người dùng chưa quen → Hỗ trợ tùy chọn “tắt” gợi ý hoặc tinh chỉnh mức độ cá nhân hóa.

3.4. Phương hướng cải tiến (sau 4 tháng)

* Marketing Campaign Targeting:
  + Tự động chọn đối tượng cho các chiến dịch marketing (email, SMS) dựa trên hành vi & phân khúc.
  + Kết nối chặt chẽ với CRM, theo dõi chỉ số chuyển đổi chiến dịch.
* Personalization of Trading Platform:
  + Gợi ý công cụ, dashboard, chỉ báo phân tích phù hợp với phong cách giao dịch của từng người (kỹ thuật, cơ bản, day-trading...).
  + Ứng dụng Reinforcement Learning để liên tục tối ưu UI/UX dựa trên tương tác thực tế.
* Mở rộng mô hình gợi ý cho chatbot, voice assistant, hỗ trợ real-time push notification.
* Áp dụng deep learning (transformer-based, seq2seq) cho việc gợi ý tin tức, nội dung nâng cao.
* Bổ sung “Social Recommendation” từ dữ liệu mạng xã hội (nếu phù hợp chính sách).