

# **Predicting Customer Churn at RetailGenius**

A graded project for AI  
Project Methodology, 2025

*Written by D. ROSSIG & H. PHAM*

## Summary

Project Strategy .....	3
Objectives .....	3
Key Performance Indicators.....	3
How does AI improve customer retention ? .....	4
Project Design .....	4
Data .....	4
Relevant data sources .....	4
Data challenges .....	4
Models .....	5
Suitable AI models .....	5
Model training, validation and testing .....	5
Model versioning and serving .....	5
Deployment.....	5
Deployment strategies.....	5
Monitoring.....	5
Monitor performance of churn prediction model over time .....	5
Plan for handling model drift and maintaining model accuracy .....	5
Project Team .....	6
Roles and expertise.....	6
Ensure cross-functional collaboration .....	7
Ensure team alignment with strategy .....	7
AI team collaboration with other departments .....	7
Project Governance & Communication .....	7
Key stakeholders for the project and communication plan.....	7
Governance instances for informing stakeholders .....	8
Communicate model insights to both technical and non-technical team .....	9
AI Project Management Methodology .....	9
Methodology .....	9
Why is Kanban suitable for our project ? .....	9
Potential risks and mitigation strategies .....	9
Handle costs and planning derivation .....	10
Mock AI project.....	10
Main sources for this report .....	11

# Project Strategy

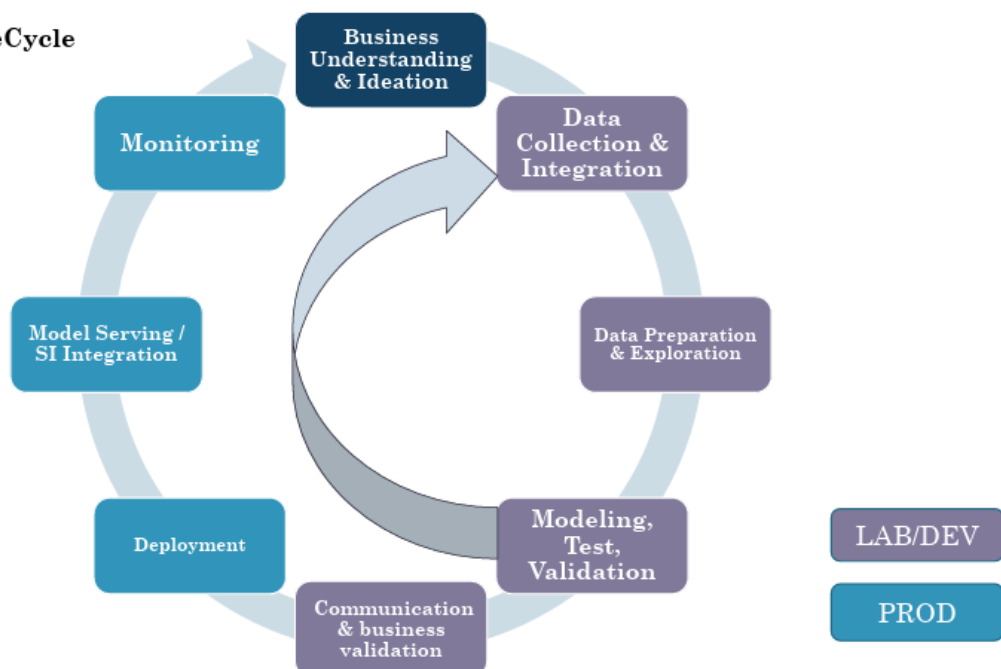
## Objectives

The primary objective is to frame if a user will stay on our e commerce platforms using churn prediction, by building and deploying a machine learning model. Anchored within the company's broader goal to enhance customer retention and sustain growth, the project is designed to deliver tangible business value through the informed application of artificial intelligence.

Related objectives are :

- Apply a multi-disciplinary approach, mixing expertise and business to solve the problem
- Produce a end-to-end AI Project Lifecycle [see scheme below]
- Enhance user experience & operational efficiency
- Maintain quality through data engineering and model performance
- Set advanced analytics and AI as foundations to RetailGenius business strategy

### AI Project LifeCycle



*AI Project LifeCycle, EPITA\_AI\_PM\_Course\_Chapter1\_2025*

## Key Performance Indicators

Let's define a thorough set of KPIs that cover both technical model performance and more general business objectives to assess the effectiveness and business impact of RetailGenius' churn prediction model. Our KPIs include :

- Product performance data
- User growth and retention (very important)
- Customer type and conversion rate
- Model improvement



## How does AI improve customer retention ?

Implementing AI at RetailGenius means identify at-risk customers before they churn, which is very important to be sure that the company continues in his upwards trend. A high churn rate not only impacts immediate revenue but also threatens long-term business growth and customer lifetime value. AI models will also analyze large and varied datasets of the company—like browsing history, purchase patterns, and customer demographics—to deliver highly personalized product recommendations and shopping experiences. Other aspects will improve because of AI : Scalable software, optimized inventory management, fraud detection...

The goal of the RetailGenius churn prediction initiative project strategy is to maintain the business's rapid customer acquisition while also greatly enhancing retention results. RetailGenius wants to predict customer churn with high accuracy by utilizing its extensive and diverse data resources, advanced data infrastructure, and knowledgeable multidisciplinary team. Strong data management procedures, scalable and safe storage options, and sophisticated ETL procedures all work together to guarantee that the project has strong organizational and technical support. RetailGenius is well-positioned to make data-driven decisions that propel long-term business success and competitive advantage in the ever-changing e-commerce landscape, thanks to its well-defined goals for increasing growth and enhancing customer retention through actionable AI insights.

## Project Design

### Data

#### Relevant data sources

Building capable and a relevant data flow means leverage what is already available in the company. RetailGenius employs a blend of SQL and NoSQL databases, cloud storage solutions, and data lakes. Structured data (such as transactions, user profiles) is typically stored in SQL databases, while unstructured or semi-structured data (like logs or clickstreams) is captured in NoSQL databases and data lakes to support scalable analysis. External data sources like Market trends are also available. Data from these diverse sources is then inputted and transformed through robust ETL (Extract, Transform, Load) pipelines, using tools such as Apache Nifi and Talend to ensure data is unified, clean, and available for the AI models.

#### Data challenges

It still poses some potential challenges, despite RetailGenius's data ecosystem being rich and varied : we need to focus ongoing attention to data quality, integration, and security. It takes lots of time for data preparation, meaning with the variety and quantity of data sources, significant time is required to handle missing data, correct inconsistencies, and engineer relevant features. Regarding storage, we need to be sure that our SQL/NoSQL systems can make integration and processes less resource intensively. Also on the secure side, we need to make sure that data processing strictly adheres to privacy regulations and internal security policies.

## Models

### Suitable AI models

The main objective of the churn prediction project at RetailGenius is to classify customers as either likely to churn or remain active, making supervised classification models the most appropriate choice. We were thinking about using either XGBoost (effective for structured/tabular data, XGBoost is well-suited to customer churn prediction because of its ability to handle missing values, prevent overfitting through regularization, and deliver high predictive accuracy) or Random Forest (robust tree-based model that excels in capturing non-linear relationships and variable interactions in customer behavior data).

### Model training, validation and testing

We would follow what we've learned at our Master to ensure the model is robust, generalizable, and actionable, the model development pipeline follows a structured and iterative process

### Model versioning and serving

All models, scripts, and configuration files are version-controlled using platforms like Git or MLFlow. Testing before serving will be done with unit tests, integration tests and business tests (after deployment). Regarding deployment, we can package our models in Docker containers to ensure portability in all environments. If we want to go the cloud route, the AWS cloud infrastructure provides enough features for serving clients.

## Deployment

### Deployment strategies

Multiple solutions are possible for deployment :

- Leverage AWS's robust infrastructure, using scalable services such as Amazon SageMaker, AWS Lambda, and containerized solutions to deploy our models
- Model is packaged independently of any single application as Data
- Use web applications enabling real-time churn risk scoring and personalized interventions
- Use internal tools, like dashboards, customer support platforms, or marketing suites

## Monitoring

### Monitor performance of churn prediction model over time

Again, there are lots of ways to monitor the churn predictions over time, but the main one is model performance metrics (Accuracy, Precision, Recall, AUC). Operations metrics like latency, uptime, resource use and Data quality metrics like data drift or input validation are also important. For those, we can use automated monitoring tools provide model accuracy and core KPIs in real-time. Alerts are configured to flag drops in performance or data drift, triggering investigation and retraining cycles as necessary

### Plan for handling model drift and maintaining model accuracy

Effectively managing model drift and maintaining sustained accuracy is critical for the long-term success of the churn prediction system at RetailGenius. Model drift is caused by business

process and customer behavior changes, putting our model at risk. For handling model drift and make sure it's maintained we can :

- Do continuous monitoring for model drift, with performance tracking, data quality checks and alert systems (MLFlow)
- Incorporate user feedback and analyze it
- Update our model with the most recent data
- Upgrade a model only if it's better
- Company-wide agreement on the accuracy standards

## Project Team

### Roles and expertise

Role	Key Responsibilities	Essential Expertise
<b>Data/AI Project Manager</b>	<ul style="list-style-type: none"><li>- Lead project lifecycle</li><li>- Coordinate technical/business teams</li><li>- Manage scope, budget, timeline</li></ul>	Project management (Agile/Scrum), business domain knowledge, stakeholder communication
<b>Data Scientist</b>	<ul style="list-style-type: none"><li>- Build and validate ML models</li><li>- Feature engineering</li><li>- Data analysis and visualization</li></ul>	Statistics, machine learning, data exploration, programming (Python, SQL), business problem solving
<b>Data Engineer / ML Engineer</b>	<ul style="list-style-type: none"><li>- Develop and maintain data pipelines</li><li>- Ensure data quality and integration</li><li>- Support deployment</li></ul>	ETL processes, big data platforms, cloud infrastructure, DevOps, data querying (SQL/NoSQL)
<b>Data Architect</b>	<ul style="list-style-type: none"><li>- Design data architecture and storage</li><li>- Ensure data governance and scalability</li></ul>	Data modeling, metadata/catalog management, database/cloud expertise, security and privacy standards
<b>Data Owner / Business Expert</b>	<ul style="list-style-type: none"><li>- Define business needs</li><li>- Ensure data relevance and quality</li><li>- Align data with objectives</li></ul>	Deep process & data knowledge, quality standards, documentation, requirement gathering
<b>Software Engineer / Application Developer</b>	<ul style="list-style-type: none"><li>- Integrate models via APIs/microservices</li><li>- Support front-end/back-end</li><li>- Ensure reliable deployment</li></ul>	Programming (Python, Java, etc.), web services, software integration, system implementation

The project's multidisciplinary team structure ensures all critical skills are represented below:

- Data and ML experts design, build, and tune predictive models.
- Data architects and engineers guarantee the reliability, scalability, and integrity of the data platform.
- Software developers and platform engineers ensure seamless integration with business systems.
- Project managers and business experts align solutions with strategic objectives and oversee efficient delivery

## Ensure cross-functional collaboration

RetailGenius can build a high-functioning, cross-disciplinary team where clear communication, shared goals, mutual respect, and agile collaboration drive the success of the churn prediction AI project and beyond, by :

- Establishing clear roles, responsibilities, and expectations
- Foster open and contextual communication
- Promote mutual understanding through workshops
- Iterate processes and solicit feedback

## Ensure team alignment with strategy

To be sure that the team aligns with the company's strategy, we need to put forward a shared way of doing things, meaning a shared strategy, building an alignment team culture and transparent communication. Good documentation like shared definitions and terminology can help as well as regularly reviewing both team culture and practices can help RetailGenius to ensure project goes well.

## AI team collaboration with other departments

The main key word here again is communication. This goes by adopting collaboration platforms such as Slack, Microsoft Teams, or Asana to centralize daily communication, task tracking, and project updates. At project kickoff, we clearly articulate the joint objectives (e.g., reducing customer churn, increasing the effectiveness of marketing campaigns) and how each department's contributions propel broader business success. We also maintain open channels for ongoing feedback, rapidly capturing frontline insights (such as common customer pain points or shifts in campaign effectiveness) to refine models and features

# Project Governance & Communication

## Key stakeholders for the project and communication plan

Below are the key stakeholders for this AI project, for us the more technical roles are important :

Title	Roles	Main tasks	Other contributions
<b>Business stakeholders</b>	Senior management, product owners, marketing executives, and customer support leaders	Set project objectives, define business priorities, make go/no-go decisions, and evaluate the impact of AI-driven insights on	Provide domain expertise, validate requirements, and measure AI's business value

		customer retention and revenue growth	
<b>Data team</b>	Data scientists, data engineers, and data analysts	Lead all stages of data preparation, model development, validation, and performance monitoring	Design and implement feature engineering, select and fine-tune models, evaluate results, and ensure data quality and integrity
<b>Technological team</b>	Data/ML architects, DevOps engineers, infrastructure specialists, software/application developers	Architect, deploy, and scale the AI solution in production environments while ensuring security, scalability, and system reliability	Oversee integration into cloud infrastructure and internal systems, maintain data pipelines, and manage ongoing technical operations

Regarding the communication plan, we can go from the “7 steps in making a communication plan” [ref] :

1. RetailGenius is a global, fast-growing e-commerce company focused on leveraging AI to minimize customer churn and enhance the user experience
2. Strengthen stakeholder alignment and buy-in for the AI churn prediction project and produce value for customer retention
3. Target audience is everyone, and especially people buying online a lot
4. At RetailGenius, we want to be transparent, collaborative, and results-oriented, emphasizing innovation and shared value. It would be amazing if the churn prediction project is essential to improving customer retention and driving company performance
5. We will use Internal, Cross-department and broad channels to communicate, for example tools like Slack, workshops or bi-weekly reports
6. The main plan is to make tech and business people work with and implement AI, meaning weekly progress updates, a big launch meeting and training sessions. The budget will partially be user as a resource for communication tools, training materials, meeting logistics, and performance tracking software. Trello for action tracking and scheduling is decent and we will do regular check-ins to update the plan and adjust as needed.
7. Analyze results through KPIs and retrospect on milestones

## Governance instances for informing stakeholders

The following governance instances and practices will be put in place to facilitate structured communication, decision-making, and monitoring across all key groups :

- Steering Committee : Provides top-level oversight, strategic guidance, and alignment with organizational goals
- Risk and Ethics Panel : Monitors ethical considerations, data privacy, and project risks related to AI deployment, to protect users
- Stakeholder Review Forums : Engage a broader group of stakeholders to gather feedback, validate deliverables, and surface new requirements



- Cross-Functional Project Board : Coordinates operational decisions and ensures seamless collaboration between business, data, technology units and even maybe new projects

## Communicate model insights to both technical and non-technical team

Regarding communicating model outputs and predictions to technical and non-technical teams, the aim is to bridge the gap between data science complexity and practical, actionable business value, facilitating widespread adoption and informed decision-making. For technical teams (data, IT...), we would communicate the detailed information about model performance, methodology and technical metrics plus add data visualizations and interactive dashboards. For non-technical teams (business stakeholders, customer support...), it would be synthetic reports, clear explanations of what the predictions mean, the impact on business KPIs, and recommend next steps. In general, we want to use synthetic reporting with story telling, for example stating the report with “Why does churn matter ?”.

## AI Project Management Methodology

### Methodology

For the AI-driven churn prediction project at RetailGenius, Kanban is an especially suitable project management methodology. Its adaptability, transparency, and focus on incremental, continuous improvement align well with the experimental and multidisciplinary nature of AI initiatives. With geographically dispersed team members and a multidisciplinary setup, deploying a digital Kanban board (e.g., Trello, Jira, Azure Boards) ensures all participants can easily monitor work status, provide input, and flag blockers in real time [ref]. With Kanban, RetailGenius ensures an agile, visible, and collaborative project environment that is ideally matched to the needs of AI-based churn prediction.

### Why is Kanban suitable for our project ?

Kanban is a highly suitable project management methodology for the RetailGenius churn prediction initiative due to its visual approach, and ability to handle the dynamic nature of AI and data science projects. Product and data science teams can effectively handle frequent changes in project priorities, requirements, and discoveries—all of which are typical in the creation of AI solutions—by using Kanban. Teams can create tools and features in a robust, dependable, and repeatable way with its visual workflow and incremental delivery structure. We also have full transparency using the Kanban “TO DO” columns. Finally, Kanban allows teams to quickly adjust priorities and adjust to new information, in contrast to Scrum, which is predicated on rigid, time-boxed sprints and predetermined commitments. This flexibility eliminates bottlenecks and rework in data science and AI development, where unforeseen obstacles, fresh insights, or changing business objectives are commonplace. This enables the team to concentrate on the most important and pressing tasks at any given time.

### Potential risks and mitigation strategies

Observing the use case, we can see some potential risks, that can be more or less mitigated :

Risk	Mitigation strategy
------	---------------------

<b>Insufficient AI/ML knowledge in the team</b>	AI and machine learning training for both technical and non-technical team members, through workshops
<b>Lack of code maintenance</b>	Adopt strict coding standards, version control, enforce code reviews, and maintain comprehensive in-line documentation
<b>Security issues, like sensitive customer data</b>	Incorporate privacy-by-design principles, role-based access controls and strict data governance across all environments
<b>Simplicity of the Model</b>	Perform thorough exploratory data analysis and feature engineering, benchmark simple models and use model explanation tools

## Handle costs and planning derivation

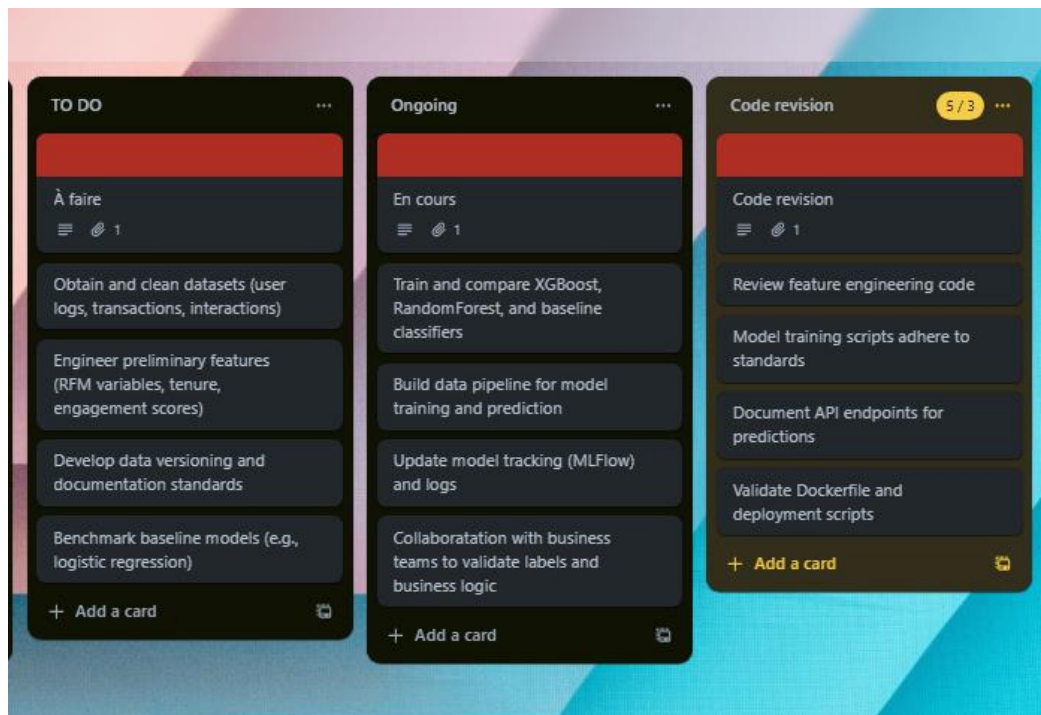
Predicting costs for labor, data storage, computer resources, and tool usage is essential for cost management and planning during iterative AI model development. These costs should then be monitored and managed at every stage of the project. Agile prioritization is supported, and overruns are avoided by employing phase gates, work-in-progress (WIP) limits, and transparent cost attribution through iteration. Teams can swiftly spot deviations and modify plans by utilizing tools for automated tracking, rapid prototyping to "fail fast," and proactive stakeholder communication. Timelines and cost estimates can be improved with ongoing reviews following each iteration, and resource allocation is further optimized by incorporating AI-driven forecasting. Despite the experimental nature of AI work, this methodical approach guarantees that the project delivers value effectively while preserving budget discipline and project momentum.

## Mock AI project

Link :

<https://trello.com/invite/b/687cbb2f00a9ac6a96a3ac42/ATTI4dc1eefd0b341ab3842d13ab007c57ec83605CC4/part-1-ai-project-mock> —

Screenshot :



## Main sources for this report

- AI Project Methodology slides, Medina Hadjem
- 7 steps of a communication plan, [OrSys](#)
- KanBan Agile method, [Medium](#)
- KanBan for AI projects, [Djaa](#)