Capstone Project Proposal

Stephen D. Gardner - MailGeniusAI

Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	The primary industry problem addressed by MailGeniusAI is the inefficiency and lack of personalization in email marketing campaigns, which leads to low engagement and high churn rates. Businesses often fail to deliver the right message at the right time to the right audience, which results in missed opportunities. MailGeniusAI aims to automate personalization, audience segmentation, and optimize email timing, leading to higher engagement and customer loyalty. Using ML/AI in this context provides a significant business benefit by improving email marketing effectiveness. By leveraging AI, businesses can boost customer satisfaction, increase repeat customer rates, and improve overall campaign performance. MailGeniusAI will offer measurable benefits, such as higher open rates, click-through rates, and conversions, contributing to long-term business success through better customer engagement and retention.
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	This problem is important to solve because personalized, timely email marketing significantly impacts customer engagement, leading to higher recurring revenue and improved market share. Customers are more likely to respond positively to well-targeted and relevant messaging, which increases customer satisfaction and loyalty. Building MailGeniusAI as an AI-powered email marketing tool, will streamline and optimize these marketing efforts, thus helping businesses maximize customer retention and drive recurring revenue. The business case for this product centers on its ability

	to enhance customer interactions, reducing churn while improving campaign effectiveness. By increasing customer happiness and personalizing experiences, MailGeniusAI will help businesses capture more market share and sustain long-term growth, making it a critical investment for competitive success.
Application of ML/AI What precise task will you use ML/AI to accomplish? What business outcome or objective will you achieve?	The problem we aim to solve with MailGeniusAI is the challenge of personalizing email marketing campaigns at scale, which is critical for increasing customer satisfaction and repeat purchases. Many businesses struggle with delivering relevant content to individual customers due to the vast amounts of data available and the need to manually tailor messages. By leveraging ML/AI, MailGeniusAI will automate the process of personalizing email content based on customer behavior, preferences, and historical data, ensuring each customer receives tailored and timely communication. This will help drive customer engagement and improve retention, directly benefiting the business by increasing recurring revenue and customer loyalty. This tool will address an industry-wide problem by offering a solution that enhances the effectiveness of email marketing through automation and data-driven decisions. By solving this problem, businesses can achieve measurable improvements in customer retention, satisfaction, and ultimately, recurring revenue streams, which are vital drivers of business success.

Success Metrics

Success Metrics What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	The success metrics for MailGeniusAI will include customer engagement rates (e.g., open and click-through rates), conversion rates, and customer retention rates. These metrics will be clearly defined and measurable through email marketing analytics tools. To establish a baseline, we will gather historical data on current email campaign performance and compare MailGeniusAI's impact over time. The baseline value will serve as a point of comparison, helping us measure how well AI-driven personalization improves overall engagement and contributes to the business goal of increasing recurring revenue and customer loyalty. We will look at clearly defined, measurable metrics to track the success of MailGeniusAI. Below are examples of business metrics we would apply to assess the effectiveness of the new AI-powered email marketing tool: 1. Open Rates: A current baseline open rate of 20% would be recorded before implementation. After deploying the product, the goal would be to improve this by 5%, targeting a 25% open rate within three months. 2. Click-Through Rates (CTR): The current baseline might show that 3% of recipients click through email links. The AI tool aims to increase personalization and relevance, driving this metric up to 4.5%. 3. Customer Retention Rates: Example - Historical data shows a 70% customer retention rate after email campaigns. By tailoring email content based on user behavior, MailGeniusAI aims to improve retention by 3%, bringing it to 73%. To establish a baseline, we would gather existing data over the past six months from the company's email marketing campaigns. This data will serve as the point of comparison to measure improvements after integrating MailGeniusAI.

Data

Data Acquisition Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?	Data for MailGeniusAI will be primarily sourced from internal email campaign performance metrics, collected through the company's Customer Relationship Management (CRM) system and email marketing platforms. This strategy reduces the need for costly external datasets. However, to strengthen the Artificial Intelligence (AI) model, anonymized external datasets on email interaction trends may be acquired. To address privacy concerns regarding Personally Identifiable Information (PII), all data will be anonymized before processing, ensuring compliance with the General Data Protection Regulation (GDPR). Data will also be continuously collected to enable real-time updates and model improvements, avoiding the limitations of a one-time batch method that requires frequent data refreshes.
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	The initial dataset for MailGeniusAI will consist of approximately 50,000 email campaigns collected over a five-year period, amounting to roughly 20 GB of data. This dataset will include metrics such as open rates, click-through rates, bounce rates, and unsubscribe rates, derived from both B2B and B2C email campaigns. Potential biases may include overrepresentation of specific industries or audience segments, such as a higher number of B2C campaigns in e-commerce compared to B2B sectors. This imbalance may skew the model's effectiveness across different industries. To mitigate bias, the data will be supplemented with additional, balanced datasets from external sources to ensure a more representative sample across industries and audience types.
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels versus any other option?	For MailGeniusAI, we decided to label the dataset using five key categories: "Open Rate," "Click-Through Rate (CTR)," "Bounce Rate," "Unsubscribe Rate," and "Campaign Type" (B2B/B2C). These labels were chosen to track core email performance metrics, as they provide critical insights into how well an email campaign is engaging its target audience. We selected these labels based on their direct correlation to user engagement and business goals such as retention and conversion. The strength of this labeling scheme lies in its focus on universally recognized email marketing metrics, which ensures consistency in analysis across different campaigns. However, one potential weakness is that the

scheme might overlook more granular behavioral data, such as time spent reading the email, which could offer deeper insights.

Possible Example Dataset:

Open Rate (%)	CTR (%)	Bounce Rate (%)	Unsubscribe Rate (%)	Campaign Type
22.5	5.3	1.1	0.3	B2B
35.8	12.7	0.9	0.5	B2C
15.6	4.1	2.2	0.8	B2B
40.3	14.6	0.5	0.1	B2C
28.9	8.3	1.4	0.6	B2B

For MailGeniusAI, the dataset was labeled with five key categories:

1. Open Rate (%): This label represents the percentage of recipients who open the email. It is a crucial indicator of the effectiveness of the subject line and the overall timing of the campaign.

2. Click-Through Rate (CTR %): CTR measures the percentage of recipients who clicked on at least one link within the email. This is a strong indicator of engagement and the quality of the email content.

3. Bounce Rate (%): This refers to the percentage of emails that were not delivered to recipients. A high bounce rate may indicate issues with the quality of the email list or technical problems with delivery.

4. Unsubscribe Rate (%): This label tracks the percentage of recipients who chose to unsubscribe after receiving the email. A high unsubscribe rate could signal that the content is not resonating with the audience or that the frequency of emails is too high.

5. Campaign Type: The label distinguishes between B2B (Business-to-Business) and B2C (Business-to-Consumer) campaigns. This classification helps to analyze performance based on the target audience.

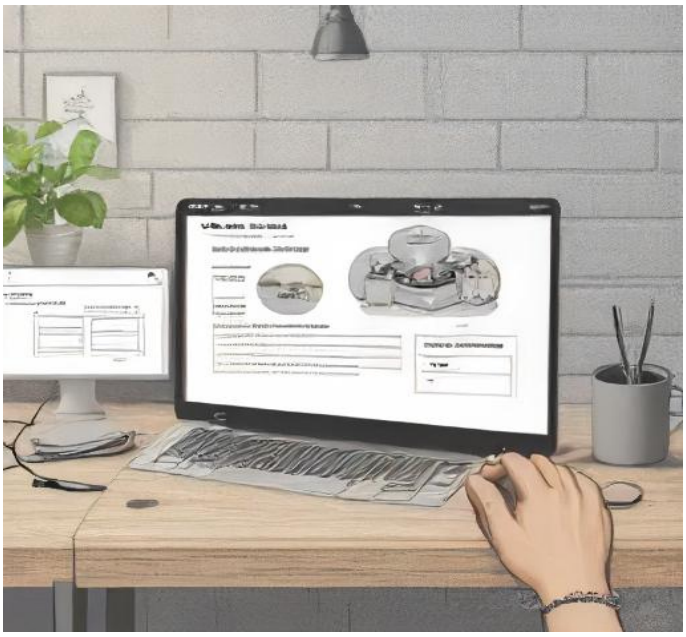
Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team, and why?	To build MailGeniusAI, the data science team will consist of several key roles. First, a Data Scientist will lead model development, focusing on algorithm selection, feature engineering, and model training. They will need strong technical skills in machine learning (ML), Python programming, and data wrangling tools like pandas and NumPy. Second, a Data Engineer will manage the data pipeline, ensuring that the model can process real-time data and handle large-scale datasets efficiently. They will require expertise in data architecture, database management, and cloud platforms like Amazon Web Services (AWS) or Microsoft Azure. Additionally, a Machine Learning Engineer will help deploy the model, ensuring it is optimized for production use, with proficiency in ML frameworks such as TensorFlow or PyTorch. On the business side, a Business Analyst will work closely with the technical team to translate business objectives into actionable data tasks. They will need strong business acumen, knowledge of customer lifecycle metrics, and an understanding of marketing strategies. Finally, a Product Manager will oversee the alignment of the AI product with business goals, ensuring that customer needs and company objectives are met. They will need excellent communication skills, project management experience, and the ability to bridge the gap between technical and non-technical teams. Together, this team will ensure that MailGeniusAI is both technically sound and aligned with the company's strategic goals.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	To measure the success of the MailGeniusAI model, key machine learning (ML) performance metrics such as accuracy, precision, recall, and the F1 score will be used. Accuracy will assess the overall correctness of email predictions, ensuring that targeted marketing emails are sent to the correct customer segments. Precision is critical for avoiding false positives, ensuring that promotional emails do not mistakenly get classified as spam, while recall will evaluate the model's ability to ensure no genuine promotional emails are missed when interacting with potential customers. The F1 score, which balances precision and recall, will help determine the model's performance in cases of imbalanced data, particularly when certain email types may dominate the dataset.

	For MailGeniusAI, a minimum F1 score of 0.90 will be required to ensure high-quality email marketing campaigns that maximize engagement without overwhelming users with irrelevant content. The success of the model will also be evaluated through ongoing A/B testing to track customer engagement rates, click-through rates, and conversion rates, ensuring that the model continues to improve over time and adapts to new data.

Minimum Viable Product (MVP)

Design What does your minimum viable product look like? Include sketches of your product.	The MailGeniusAI Minimum Viable Product (MVP) will consist of a streamlined user interface that allows businesses to upload customer email lists, create personalized email templates, and track real-time engagement metrics such as open rates and click-through rates. The MVP design will feature three key components: the Dashboard for performance metrics, the Email Campaign Builder for creating and personalizing email content, and the Analytics Page for viewing campaign effectiveness over time. Initial sketches will focus on simplicity and functionality, providing users with an intuitive interface that allows them to create email campaigns in minutes. The MVP will include basic A/B testing capabilities to help users optimize their email strategies based on early feedback. The rollout will prioritize ease of integration with popular customer relationship management (CRM) platforms like HubSpot or Salesforce to ensure seamless adoption for businesses. Below is a rough MailGeniusAI sketch:

	 Source: https://deepai.org/machine-learning-model/text2img
Use Cases What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?	The MailGeniusAI product is designed for small-to-medium-sized business owners, marketing managers, and email campaign specialists who rely on email marketing to drive customer engagement and revenue growth. One prototypical user might be a Marketing Manager at an e-commerce business looking to automate personalized email campaigns to segmented customer lists. Another persona could be a Business Owner seeking an easy-to-use platform that integrates with existing CRM tools to increase repeat customers. Epic-level use cases include automating the creation of personalized email campaigns, A/B testing subject lines to improve open rates, and analyzing campaign performance to refine future strategies. Users will access the product via a web-based interface, designed for ease of use on both desktop and mobile, with API integrations to pull data from platforms like HubSpot, Salesforce, and Google Analytics.
Roll-out	The go-to-market plan for MailGeniusAI will focus on a phased rollout. The pre-launch phase will

How will this be adopted? What does the go-to-market plan look like?	involve a beta program targeting small business owners and marketing teams in key industries, collecting feedback to refine the product. This phase will also include a marketing campaign to generate awareness, leveraging social media, email marketing, and partnerships with CRM platforms. The launch phase will consist of a broader public release, with a focus on onboarding users through webinars, tutorials, and customer support to ensure a smooth adoption. The post-launch plan will involve continuous monitoring of user feedback, rapid iteration based on insights, and expanding features based on early user data. Major milestones include product development completion, beta release, marketing campaign, public launch, and post-launch user engagement.

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	After the Minimum Viable Product (MVP) launch of MailGeniusAI, the product will be iteratively improved by incorporating real-world data into the model. Real-world data may differ from training data in various ways, such as changes in customer behavior or new email marketing trends. To address these variations, the model will use active learning to continuously refine its predictions by incorporating user feedback and performance metrics. Additionally, A/B testing will be employed to compare different product features and adjustments to optimize for performance, ensuring the product stays relevant and adapts to changing user needs over time.
Monitor Bias	To monitor and mitigate unwanted bias in MailGeniusAI,

How do you plan to monitor or mitigate unwanted bias in your model?	we will employ bias-aware algorithms such as re-weighting and resampling. These algorithms help ensure that different demographic groups are represented fairly in the training data. Re-weighting adjusts the importance of certain data points based on underrepresented groups, while resampling balances the dataset by duplicating or removing samples to ensure even distribution. Post-launch, we will also implement fairness metrics and conduct regular audits using these algorithms to continuously assess and reduce any emerging biases.