



DataHack

The word "DataHack" is displayed in a large, bold, white sans-serif font. It is overlaid on a background of a dense grid of various firearms, including handguns, rifles, and machine guns, all rendered in a dark gray color. The text has a slight shadow effect and is partially obscured by the gun icons.

missions

The word "missions" is written in a large, bold, white sans-serif font with a black outline. It is positioned below the "DataHack" text and also sits atop the same grid of firearms. The letters have a slight shadow and are partially hidden by the gun icons.

DS | AI | MLOPS | NLP

The text "DS | AI | MLOPS | NLP" is centered at the bottom of the image in a large, bold, white sans-serif font. It is placed directly over the bottom row of the gun grid.

MISSIONS

PS1 - Data Science

Problem Statement:

Flight Profile Generation for Enhanced Aviation Operations

Objective:

Liberty City's aviation system is in disarray due to rampant criminal activity, disrupting flights and schedules. Participants are tasked with developing a dynamic flight profile generation algorithm using the FlightQuest dataset to restore order. The algorithm will aggregate real-time weather data, flight statuses, and historical performance to mitigate delays and improve decision-making for pilots and emergency teams. It will provide critical insights to manage disruptions caused by weather and streamline operations. A user-friendly dashboard will visualize key metrics to enhance aviation efficiency and safety across Liberty City.

Key Features:

1. Dynamic Flight Profile Dashboard -

- Develop an algorithm to generate dynamic flight profiles, incorporating flight status, current weather conditions, historical performance, and estimated fuel consumption.
- Create an interactive dashboard to showcase these profiles, including visualizations of key metrics like average delays, weather impacts, and fuel consumption patterns.

2 .ETA Prediction and Delay Analysis -

- Models that predict flight delays based on a variety of factors, including historical and current weather data, air traffic conditions, and operational issues.
- Real-time expected arrival time (ETA) prediction models that adjust based on in-flight updates and changing circumstances, integrating both current and historical data to provide accurate time estimates.

3. Weather Impact Assessment Models -

- A safety hazard identification model that assesses severe weather conditions and their potential impact on flight safety and operations.
- Predictive analytics for understanding how weather variations affect flight schedules, delays, and overall operational efficiency.

4. Interactive Data Visualization -

- A user-friendly, deployed web interface for interactive exploratory data analysis (EDA), allowing users to visualize flight times, durations, weather conditions, flight statuses, fuel consumption, and costs.
- Interactive graphs and bar charts that users can customize based on their preferences to better understand key metrics and trends.



Bonus Points:

- Customized Alerts for weather conditions and flight statuses
- Cost Estimation



MISSIONS

PS2- AI

Problem Statement:

Adaptive Quiz Platform Leveraging MultiModal Flashcards

Objective:

Step into the high-energy environment of Los Santos by designing and developing an AI-driven adaptive quiz platform tailored for elite operatives. This system generates multimodal flashcards—including text, images, videos, and audio—to optimize personalized training experiences. Leveraging advanced AI/ML techniques, the platform intelligently adjusts the difficulty of flashcards and quizzes in real-time based on operatives' performance, response times, and feedback, ensuring they remain mission-ready and capable of handling any challenge that comes their way.

Key Features:

1. Personalized Adaptive Multimodal Flashcards:

- Automatic Generation: Create flashcards in various formats—including text, images, videos, and audio—tailored to educational topics.
- Dynamic Difficulty Adjustment: Employ reinforcement learning to monitor user performance in real-time, adjusting the difficulty of flashcards based on accuracy, response speed, and the number of attempts to ensure optimal learning retention.
- Customization: Adapt flashcard content based on user input and specific curriculum goals to meet individual learning needs.

2. Performance Insights & Recommendations:

- Provide personalized insights and study recommendations by analyzing historical performance data.
- Help users focus on weak areas and improve over time.

3. Unstructured Inputs:

- Allow users to dynamically generate flashcards from unstructured inputs such as scanned PDFs, documents, PPTs, or images.

Brownie Points:

- Consider incorporating features such as knowledge graph integrations, xAI, Spaced Repetition, Gamification, Personalized Path generation using segmentation, Material Suggestions

MISSIONS

PS3 - MLOps

Problem Statement:

MLOps Pipeline for Paleo Bay Restaurant Network

Objective:

In the picturesque and thriving coastal town of Paleo Bay, you are tasked with designing and developing an MLOps pipeline to process restaurant data, transforming raw sales and inventory information into actionable insights. Each restaurant operates independently with unique data structures and schemas, requiring a flexible and scalable solution to ensure seamless integration and analysis across the network.

Key Features:

- **Automated Data Ingestion and Transformation:**
Build an MLOps pipeline that automates the ingestion and transformation of raw restaurant data and master data tables, accommodating the varying structures and schemas of each establishment.
- **Intelligent Issue Detection and Recommendations:**
Develop a suggestion model that flags critical issues such as suspicious transactions, malpractices, theft, stock depletion, and low sales. Provide improvement recommendations using a clear rating scale (e.g., 1-5 or color-coded from red to green) to indicate severity.
- **Multi-Calendar Predictions:** Implement predictive models that generate forecasts based on different calendar systems (Gregorian, Hindu, Islamic, etc.), catering to the diverse operational timelines of restaurants
- **Continuous Feedback Loop:** Create a feedback mechanism where users can rate the usefulness of the suggestions, feeding this data back into the model to enhance future predictions and ensure continuous improvement.



Brownie Points:

Consider analyzing historical data and time series forecasting, automated dashboards based on prompts



MISSIONS

PS4 - NLP

Cybersecurity Platform Featuring an AI-Driven Assessment Bot

Objective:

Welcome to the gritty underbelly of Liberty City, where every organization is just one step away from a devastating cyberattack. Your mission: develop an AI-driven cybersecurity assessment bot that will navigate the dark alleys of a company's digital infrastructure, probing for weaknesses with 300-400 dynamic, relentless questions across domains like Network Security, Data Protection, and Incident Response. This bot is more than just a tool—it's an elite operator, adapting to every answer, calculating real-time risk scores, and delivering hard-hitting reports with actionable insights. In a city where the next attack is always looming, your platform will be the difference between survival and digital chaos.

MISSIONS

Develop a web-based platform with an AI-powered bot that conducts a comprehensive cybersecurity assessment through 300-400 dynamic questions covering the following domains:

- Network Security
- Data Protection
- Incident Response
- Compliance

The bot will engage with the company's Single Point of Contact (SPOC) to gather information, and based on their responses, will calculate a risk score that represents the organization's overall cybersecurity posture. The bot will also generate a detailed report with actionable insights and recommendations, helping the company proactively strengthen its cybersecurity defenses.

missions

Key Features:

1. Dynamic and Context-Aware Questioning:
 - Feature: The bot will use Natural Language Processing (NLP) to ask dynamic and relevant questions based on the company's previous responses. Follow-up questions will be triggered if vulnerabilities or areas of concern are identified.
 - Example: If the organization lacks multi-factor authentication (MFA), the bot will follow up with questions regarding other access control mechanisms. This ensures a comprehensive assessment of each domain.

Note: The use of open-source LLMs or fine-tuning small language models (SLMs), which are more secure for privacy reasons and require less computational power is crucial.

MISSIONS

2. Assessment of Infrastructure Documentation:

- Feature: Instead of starting with direct questions, the bot will first allow the company to upload their infrastructure documentation (e.g., network diagrams, security policy documents, configuration files). The bot will use NLP and text analysis techniques to analyze this documentation, identifying potential vulnerabilities, gaps, or non-compliance areas. Based on this analysis, the bot will generate follow-up questions to probe further into specific areas that need clarification or additional detail.
- Example: If the bot detects that there is no mention of encryption protocols in the uploaded network configuration, it will ask follow-up questions about how the company handles data encryption, probing deeper into areas where information is lacking.

Tech Suggestion: Use OCR (if needed) to convert scanned documents into machine-readable text and apply entity extraction techniques to identify key components like firewalls, encryption protocols, and access controls.

MISSIONS

3. Real-Time Risk Scoring:

- Feature: The bot will use machine learning models to analyze responses and generate a real-time risk score. This score provides an instant snapshot of the organization's security posture and highlights areas requiring attention.
- Example: If incident response controls are found to be weak, the risk score in that domain will increase, marking it as a high-priority area for improvement.

4. Detailed Report and Recommendations:

- Feature: After completing the assessment, the bot will generate a comprehensive report that includes actionable recommendations. The report will prioritize high-risk areas and suggest tailored security improvements.
- Example: The bot may recommend specific encryption tools for data protection or advise scheduling regular security audits for areas marked as high-risk.

MISSIONS

5. Basic Compliance Check:

- Feature: The bot will verify the organization's practices against common standards like GDPR and HIPAA, flagging areas of non-compliance and providing suggestions for remediation.
- Example: If certain GDPR requirements are not met, the bot will flag this and suggest corrective actions, such as implementing stronger data access controls or regular data protection assessments.

Brownie Points:

- Report Generation from scanned incident reports
- Realtime updates from public datasets such as those tracking credential leaks
- Contextual Learning with Knowledge Graphs
- Topic Modelling for Incident Reports