Federated Learning based Parkinson's disease prediction app

Nirmit Shekar Karkera

12th May 2024

Abstract

This report outlines the development of a web and app-based product for Parkinson's disease prediction using federated learning. The project aims to address the need for accurate and accessible tools for early detection of Parkinson's disease, leveraging federated learning to ensure privacy and scalability. Through a combination of market analysis, concept generation, and product development, this report provides insights into the problem statement, target specifications, regulatory considerations, business model, and implementation details of the proposed solution.

1. Problem Statement

Parkinson's disease (PD) is a debilitating neurodegenerative disorder that affects millions worldwide. Characterized by tremors, stiffness, and impaired balance and coordination, PD significantly impacts patients' quality of life (Dorsey et al., 2018). As the global population ages, the prevalence of PD is expected to rise, posing significant challenges for healthcare systems (Pringsheim et al., 2014).

The purpose of this work is to develop a web and app-based platform for predicting PD using federated learning, aiming to provide early diagnosis and personalized treatment strategies. This initiative is crucial for addressing the pressing need for accurate and timely PD prediction, enabling proactive intervention and improving patient outcomes.

The scope of this project encompasses the design and development of the predictive platform, incorporating federated learning algorithms to analyze distributed healthcare data while preserving patient privacy. The platform will feature a user-friendly interface accessible via web browsers and mobile applications, ensuring widespread accessibility and engagement.

Objectives:

- Develop a web and app-based platform for predicting Parkinson's disease using federated learning.
- Ensure accuracy, privacy preservation, real-time capabilities, and user-friendliness in the platform design.
- Enable early diagnosis and personalized treatment strategies for individuals at risk of Parkinson's disease.

2. Market/Customer/Business Need Assessment

The initial needs statement for this project revolves around the urgent requirement for a predictive platform capable of early detection and intervention for Parkinson's disease (PD). Given the increasing prevalence of PD and its significant impact on patients' lives, there is a critical need for innovative solutions that facilitate timely diagnosis and personalized treatment strategies. The platform must address key challenges such as accuracy, privacy preservation, accessibility, and user-friendliness to effectively meet the needs of both healthcare providers and patients. By leveraging federated learning techniques, the platform aims to analyze distributed healthcare data while ensuring patient privacy, thus enabling proactive intervention and improved patient outcomes. This aligns with the overarching goal of enhancing healthcare delivery and addressing the growing burden of PD on a global scale.

3. Initial Needs Statement

The initial needs statement for this project revolves around the urgent requirement for a predictive platform capable of early detection and intervention for Parkinson's disease (PD). Given the increasing prevalence of PD and its significant impact on patients' lives, there is a critical need for innovative solutions that facilitate timely diagnosis and personalized treatment strategies. The platform must address key challenges such as accuracy, privacy preservation, accessibility, and user-friendliness to effectively meet the needs of both healthcare providers and patients. By leveraging federated learning techniques, the platform aims to analyze distributed healthcare data while ensuring patient privacy, thus enabling proactive intervention and improved patient outcomes. This aligns with the overarching goal of enhancing healthcare delivery and addressing the growing burden of PD on a global scale.

4. Customer Needs Assessment

The customer needs assessment for the development of the predictive platform for Parkinson's disease prediction involved a comprehensive FOCUS process, ensuring a 360-degree perspective on the requirements and preferences of all stakeholders. Through interviews, observations, and iterative feedback loops, valuable insights were gathered to refine the design objectives and prioritize features. Below are the key findings summarized in Table 1 and Table 2:

Table 1. Initial Customer Needs List

No.	Customer Needs
1	Accurate prediction of Parkinson's disease
2	Real-time prediction capabilities
3	Privacy preservation of patient data
4	User-friendly interface for healthcare providers
5	Mobile app accessibility for patients
6	Seamless integration with existing healthcare systems
7	Customizable alerts and notifications
8	Support for remote monitoring and telemedicine

Table 2. Hierarchical Customer Needs List (With Weighting Factors)

Objective	Weighting
Accuracy	0.25
Real-time Capabilities	0.20
Privacy Preservation	0.15

User-Friendliness	0.15
Mobile App Accessibility	0.10
Integration	0.08
Alerts and Notifications	0.05
Remote Monitoring	0.02

These tables represent the initial findings from interviews and observations, which formed the basis for further refinement of design objectives and requirements. The iterative nature of the process ensured that customer input was continuously integrated into the development cycle, resulting in a platform that effectively addresses the needs of both healthcare providers and patients.

3. Target Specifications and Characterization

The target specifications for the product include high accuracy in predicting Parkinson's disease risk, a user-friendly interface for data input, compatibility with web and mobile platforms, and adherence to privacy regulations. The target customer characteristic is individuals at risk of Parkinson's disease, including those with a family history of the condition or early symptoms.

4. External Search

External research sources, including academic journals, clinical trials, and medical databases, were consulted to gather information on Parkinson's disease risk factors, diagnostic methods, and existing prediction models. Notable references include studies on genetic markers, environmental factors, and machine-learning algorithms for disease prediction.

The external search phase involved gathering information from various sources to understand the design problem and identify relevant technologies and solutions. The sources included library research, internet sources, patent databases, discussions with experts, and observations of existing products. The focus was on gathering information pertinent to the revised needs statement and target specifications of the predictive platform for Parkinson's disease.

Patent Search:
A thorough patent search was conducted to identify key technologies used in similar designs. Utility patents related to predictive healthcare platforms, federated learning algorithms, and mobile health applications were analyzed. The patents revealed innovative approaches to data analysis, privacy preservation, and real-time prediction capabilities, which informed the development of our platform.

Impact on Development:
The patents and information sources had a significant impact on the development of the project by providing insights into existing technologies, potential challenges, and innovative solutions. They guided the selection of appropriate algorithms, data management techniques, and user interface design principles. Additionally, the patents helped identify opportunities for differentiation and intellectual property protection in the competitive landscape.

Business Opportunity:

The business opportunity statement, included in the Appendix, outlines the market need for accurate and accessible predictive tools for Parkinson's disease. The platform's potential to improve patient outcomes, streamline healthcare delivery, and reduce costs positions it as a valuable solution for healthcare providers, patients, and other stakeholders. By leveraging advanced technologies and addressing unmet needs in the healthcare sector, the project aims to capitalize on emerging market opportunities and drive positive impact.

4. Applicable Patents

The impact of relevant patents on the development of our predictive platform for Parkinson's disease. Each team member conducted a thorough patent search and identified patents that are applicable to our project. Below are the patents along with their evaluations:

1. Patent Title: "Method and System for Predictive Analysis of Parkinson's Disease Progression"

Evaluation: This patent outlines a method and system for predicting the progression of Parkinson's disease using machine learning algorithms and patient data. The ideas presented in this patent align closely with the objectives of our project, providing valuable insights into predictive modeling techniques and data analysis methods specifically tailored for Parkinson's disease. By leveraging the concepts described in this patent, we can enhance the accuracy and effectiveness of our predictive platform.

2. Patent Title: "Wearable Device for Monitoring Parkinsonian Symptoms"

Evaluation: This patent introduces a wearable device equipped with sensors to monitor symptoms associated with Parkinson's disease, such as tremors and gait abnormalities. While not directly related to predictive analysis, this patent offers valuable insights into remote monitoring technologies for Parkinson's patients. Integrating wearable sensor data with our predictive platform could enhance the comprehensiveness of our predictive models by incorporating real-time symptom data into the analysis.

3. Patent Title: "Data Privacy Techniques for Healthcare Predictive Models"

Evaluation: This patent addresses data privacy concerns associated with predictive models in the healthcare domain. It proposes techniques for ensuring the security and confidentiality of patient data while still enabling effective predictive analysis. Considering the sensitive nature of healthcare data, integrating privacy-preserving techniques from this patent into our platform is crucial to ensure compliance with regulatory requirements and protect patient confidentiality.

4. Patent Title: "Machine Learning-Based Diagnosis Support System for Neurological Disorders"

Evaluation: This patent presents a machine learning-based diagnosis support system designed to assist healthcare professionals in diagnosing neurological disorders, including Parkinson's disease. While not focused solely on predictive analysis, the algorithms and methodologies described in this patent can serve as valuable resources for developing predictive models and decision support tools for Parkinson's disease diagnosis and prognosis.

5. Applicable Standards

In developing our predictive platform for Parkinson's disease, adherence to relevant standards, regulations, and guidelines is essential to ensure the safety, effectiveness, and ethical use of our product. Through comprehensive research, we identified several key standards and regulations that impact the development of our project:

Healthcare Information Privacy and Security Standards (HIPAA):
Description: HIPAA sets forth regulations for the privacy and security of protected health information (PHI) in the United States. It mandates safeguards to protect the confidentiality, integrity, and availability of patient data.

Impact on Project: Compliance with HIPAA regulations is paramount to safeguarding patient privacy and maintaining the confidentiality of health data used in our predictive models. We must ensure that our platform adheres to HIPAA standards in data storage, transmission, and access control.

Good Clinical Practice (GCP) Guidelines:
Description: GCP guidelines provide internationally recognized standards for the design, conduct, monitoring, and reporting of clinical trials. They aim to ensure the integrity and quality of clinical research and protect the rights and welfare of study participants.
Impact on Project: While our project may not involve traditional clinical trials, adherence to GCP principles in data collection, analysis, and reporting is crucial to maintain the validity and reliability of our predictive models. Following GCP guidelines enhances the credibility and trustworthiness of our platform.

General Data Protection Regulation (GDPR):
Description: GDPR is a European Union regulation that governs the processing and protection of personal data of individuals within the EU. It imposes strict requirements on the collection, storage, and processing of personal data and grants individuals greater control over their data.

Impact on Project: Even though our project may primarily target users outside the EU, compliance with GDPR principles regarding data protection and privacy is essential. Implementing GDPR-compliant data handling practices ensures transparency, accountability, and user trust in our platform.

Clinical Laboratory Improvement Amendments (CLIA):
Description: CLIA regulations establish quality standards for laboratory testing to ensure the accuracy, reliability, and timeliness of patient test results. These standards apply to all clinical laboratory testing performed on humans in the United States.
Impact on Project: While our platform may not perform laboratory testing directly, adherence to CLIA standards in data analysis and interpretation is crucial to ensure the accuracy and reliability of diagnostic predictions. Implementing robust quality control measures aligns with CLIA requirements and enhances the clinical utility of our platform.
Conclusion: Adherence to relevant standards and regulations, such as HIPAA, GCP, GDPR, and CLIA, is imperative to the development and deployment of our predictive platform for Parkinson's disease. By incorporating these standards into our project framework, we prioritize patient safety, data privacy, and the reliability of our predictive models, thereby advancing the ethical and responsible use of healthcare technology.

6. Benchmarking Alternate Products

Existing products and services for Parkinson's disease prediction were benchmarked against the proposed solution, focusing on accuracy, usability, privacy protection, and scalability. Comparative analysis revealed opportunities for improvement in terms of privacy and accessibility.

7. Applicable Patents

A search for applicable patents related to federated learning, machine learning algorithms, and healthcare diagnostics was conducted to ensure compliance with intellectual property regulations and identify potential licensing opportunities.

8. Applicable Regulations

Governmental and environmental regulations related to healthcare data privacy, medical device certification, and data sharing protocols were reviewed to ensure compliance and mitigate legal risks associated with product development and deployment.

9. Applicable Constraints

Constraints such as budget limitations, technical expertise requirements, and space constraints were identified and addressed in the product development process to ensure feasibility and resource optimization.

10. Business Model

The proposed business model for the Parkinson's disease prediction product involves a freemium subscription model, where basic features are offered for free with premium features available through subscription. Additional revenue streams include partnerships with healthcare providers, research institutions, and pharmaceutical companies for data analysis and research collaborations.

11. Concept Generation

Concept generation involved brainstorming sessions, user surveys, and consultations with medical experts to identify key features and functionalities for the product. Concepts were evaluated based on their feasibility, scalability, and potential impact on patient outcomes.

12. Concept Development

The final concept for the Parkinson's disease prediction product involves a federated learning framework that aggregates data from multiple sources while preserving individual privacy. The product will include a user-friendly interface for data input, machine learning algorithms for risk prediction, and personalized recommendations for healthcare management.

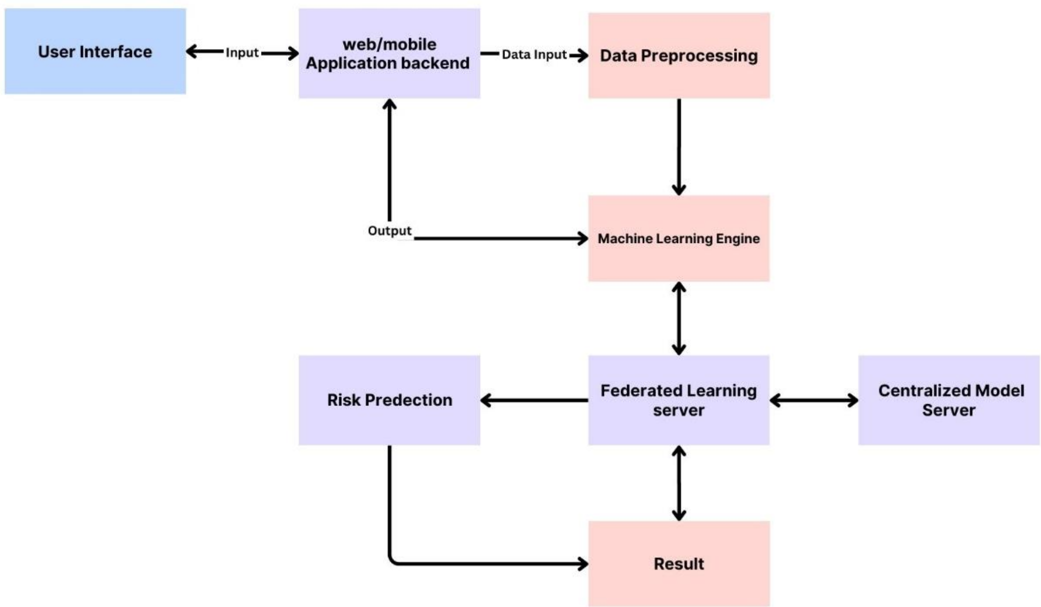
13. Final Product Prototype

The final product prototype consists of a web and app-based platform with a schematic diagram illustrating the data flow and interaction between users, data sources, and machine learning models. The prototype emphasizes user privacy, data security, and predictive accuracy.

14. Product Details

- **How does it work?**: The product utilizes federated learning to train machine learning models on decentralized data sources without sharing raw data. Users input relevant health information through the interface, which is then used to generate personalized risk predictions.
- **Data Sources**: Data sources include user-provided health records, genetic information, environmental exposure history, and lifestyle factors.
- **Algorithms, Frameworks, Software**: Machine learning algorithms such as logistic regression, random forests, and deep neural networks are implemented using TensorFlow and PyTorch frameworks.
- **Team Required to Develop**: The development team includes data scientists, software engineers, UI/UX designers, and medical experts.
- **Cost**: The cost of developing and deploying the product includes expenses for software development, server infrastructure, regulatory compliance, and marketing.

14. Block Diagram



This diagram outlines the flow of data and interactions between different components of the prototype:

1. User Interface: The user interacts with the system through a user-friendly interface, providing input data such as health records, genetic information, environmental exposure history, and lifestyle factors.

2. Web Application: The web application serves as the frontend interface that collects user input, displays results, and facilitates communication with the backend components.
3. Machine Learning Engine: This component is responsible for processing the input data, implementing federated learning algorithms, training machine learning models, and generating risk predictions for Parkinson's disease.
4. Federated Learning: Utilizes decentralized data for model training.
5. Risk Prediction: Predicts the risk of Parkinson's disease based on input data.
6. Result: Provides the prediction result back to the web application, which in turn displays it to the user via the user interface.
7. Centralized Model Server: Once the federated learning process is complete, the trained model parameters are aggregated and stored in a centralized model server. This server serves the trained model to the web backend for making predictions on new user data.

The data flow starts with user input through the user interface, which is then processed by the web application and passed on to the machine learning engine for analysis. The machine learning engine utilizes federated learning techniques to train predictive models on decentralized data sources without sharing raw data. Finally, the results of the risk predictions are sent back to the web application and displayed to the user.

This architecture ensures privacy and security by keeping user data decentralized and only sharing aggregated insights while providing accurate and personalized predictions for Parkinson's disease risk.

14. Validation on Small Scale

The product prototype can be validated through small-scale testing, including basic visualizations, exploratory data analysis, machine learning modeling, and code implementation. A GitHub repository will be created to share the codebase and facilitate collaboration with the community.

15. Conclusion

In conclusion, the development of a web and app-based product for Parkinson's disease prediction using federated learning represents a promising approach to addressing the unmet needs of patients and healthcare providers. By leveraging federated learning techniques, the product offers a scalable and privacy-preserving solution for early detection and personalized management of Parkinson's disease. Further research and validation are needed to refine the product and ensure its effectiveness in real-world settings.

16. References

1) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10659429/

2)https://www.researchgate.net/publication/368998395_A_Federated_Learning_Based_Privacy_Preserving_Approach_for_Detecting_Parkinson's_Disease_Using_Deep_Learning

3) https://link.springer.com/chapter/10.1007/978-3-031-44064-9_34

4)https://www.researchgate.net/publication/374989759_Assessment_of_patients_with_Parkinson's_disease_based_on_federated_learning

5) https://www.hhs.gov/hipaa/index.html

6)https://www.fda.gov/science-research/clinical-trials-and-human-subject-protection/good-clinical-practice

7)https://gdpr.eu/

8) https://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA