BANGALORE INSTITUTE OF TECHNOLOGY
K.R Road, V.V Puram, Bengaluru-04
DEPARTMENT OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

“Privacy Protected and Federated Adaptive EdTech Learning Platform”

Presented by

Kollipara Sai Sandeep	1BI20AI022
Shreyas R S	1BI20AI046
Somula Jaswanth Reddy	1BI20AI048
Vivek Pandith D V	1BI20AI056

VII Semester

Introduction

- Federated learning is a machine learning technique that trains an algorithm via multiple independent sessions, each using its own dataset. This approach stands in contrast to traditional centralized machine learning techniques where local datasets are merged into one training session, as well as to approaches that assume that local data samples are identically distributed.
- Traditional educational systems often adopt a one-size-fits-all approach, treating students as uniform entities rather than recognizing their unique strengths, weaknesses, and learning preferences.
- The "Privacy Protected and Federated Adaptive Learning EdTech Platform" aims to address these pressing challenges. By leveraging cutting-edge technologies, we endeavor to offer a learning environment that is dynamic and responsive to individual student performance.

Objectives

Domain Related Objectives

- To implement robust privacy-preserving techniques within a framework to solve Privacy challenge in Federated Learning.
- Distributed model training across edge devices to reduce the need for central data training, minimizing privacy and security risks.

Application Related Objectives

- To build a robust and secure infrastructure for collecting, storing, and processing student data, ensuring data integrity and confidentiality.
- Utilize federated learning to adapt educational content recommendations and assessments to individual students' abilities and preferences.
- Create a user-friendly interface that simplifies interaction with the federated learning platform, making it accessible to students and administrators.

Literature Survey

- **An overview on federated learning**
  - Authors - Chen Zhang , Yu Xie , Hang Bai , Bin Yu , Weihong Li, Yuan Gao
  - Year Published - 2021
  - Proposed idea - This paper systematically introduces the existing works of federated learning from five aspects: data partitioning, privacy mechanism, machine learning model, communication architecture and systems heterogeneity. Summarizes the characteristics of existing federated learning, and analyzes the current practical applications of federated learning.
  - Main drawback is communication overhead primarily caused by the need for data exchange and model updates between the local devices (clients) and the central aggregator (server).

- **Federated Learning Challenges, methods, and future directions**
  - Authors - Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith
  - Year Published - 2020
  - Proposed idea - Decentralized machine learning approach where models are trained locally on individual devices or servers, preserving data privacy. Model updates, not raw data, are shared and aggregated to create a global model, making it ideal for privacy-sensitive applications. It allows collaborative model training without centralizing data.
  - Core challenges in federated learning are expensive communication, systems heterogeneity, statistical heterogeneity and sometimes privacy is provided at the cost of reduced model performance or system efficiency

- **Privacy Preservation in Federated Learning**
  - Authors - Nguyen Truong, Kai Sun, Siyao Wang, Florian Guitton, Yike Guo
  - Year Published - 2021
  - Proposed idea - This paper aims at addressing the critical challenge of data privacy and security in the context of AI and machine learning applications by exploring a systematic survey of state-of-the-art privacy-preserving techniques that can be employed in Federated Learning.
  - Main drawback is that it may not address the legal and regulatory requirements of data privacy, which can vary across regions and industries, causing potential challenges in achieving full GDPR compliance for FL-based systems.

- **A Hybrid Approach to Privacy-Preserving Federated Learning**
  - Authors - Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R. and Zhou, Y
  - Year Published - 2019
  - Proposed idea - Proposed an idea to perform Federated Learning that combines Differential Privacy and Secure Multiparty Computation to improve model accuracy while preserving provable privacy guarantees and protecting against extraction attacks and collusion threat.
  - Differential Privacy alone cannot provide complete privacy and hence on combining with Secure multipart Computation additionally guarantees any messages exchanged without DP protection are not revealed and therefore do not leak any private information.

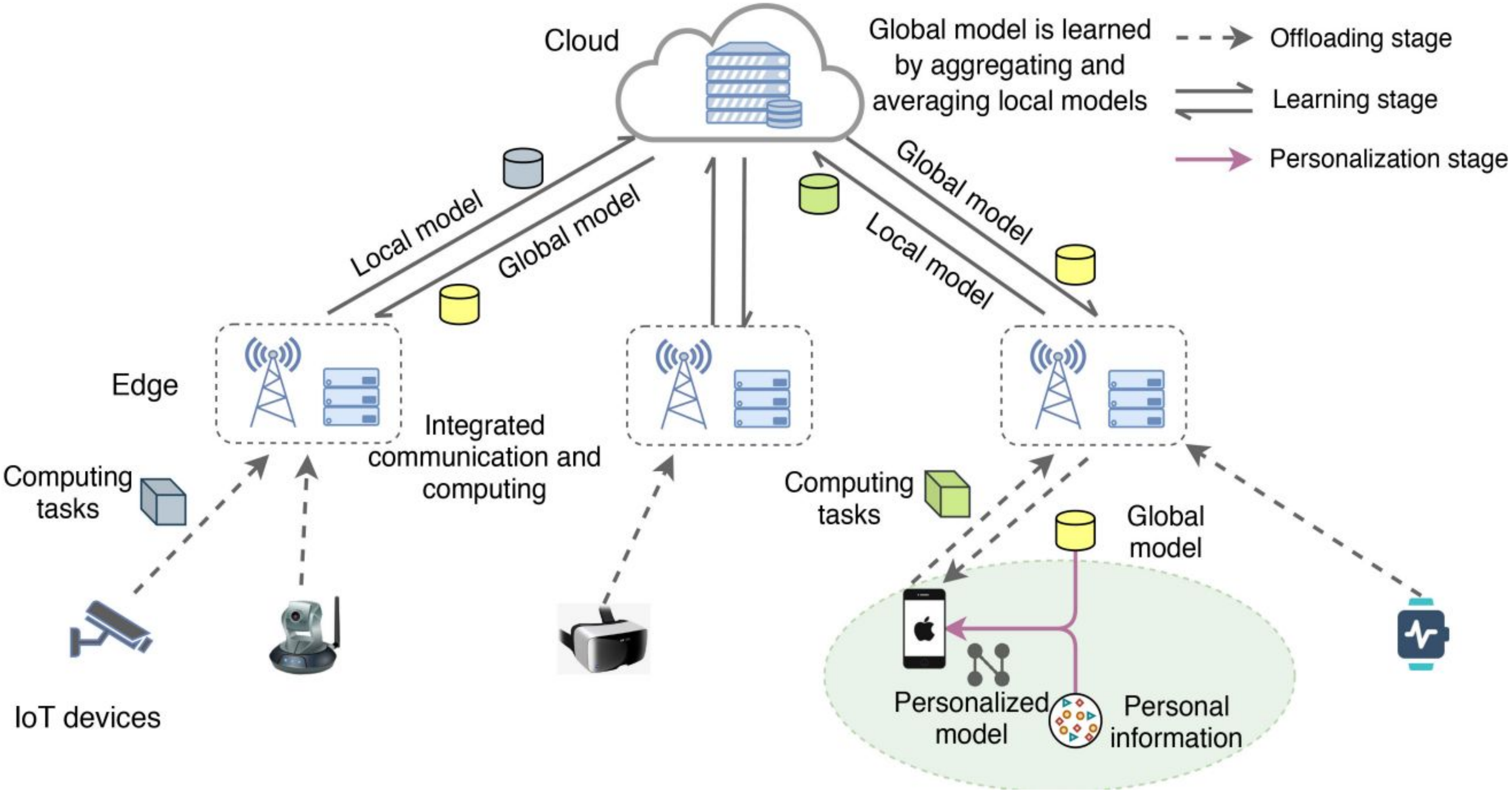
- **FedML-HE: An Efficient Homomorphic-Encryption-Based Privacy-Preserving Federated Learning System**
  - Authors - Weizhao Jin, Yuhang Yao, Shanshan Han, Carlee Joe-Wong, Srivatsan Ravi, Salman Avestimehr, Chaoyang He
  - Year Published - 2022
  - Proposed idea - FedML-HE, An efficient HE-based secure federated aggregation that provides a user/device-friendly deployment platform. This uses a novel overhead optimization scheme, significantly reducing both computation and communication overheads during deployment while providing customizable privacy guarantees.
  - Main drawback is the Performance trade-off, which comes with the usage of homomorphic encryption. Depending on the hardware resources on user devices, these trade-offs may limit the practicality and speed of the system.

- **Privacy-Preserving Federated Learning Using Homomorphic Encryption**
  - Authors - Jaehyoung Park and Hyuk Lim
  - Year Published - 2022
  - Proposed idea - Using the Homomorphic Encryption scheme, the proposed privacy-preserving federated learning (PPFL) algorithm enables the centralized server to aggregate encrypted local model parameters without decryption. Furthermore, the proposed algorithm allows each node to use a different HE private key in the same FL-based system using a distributed cryptosystem.
  - Main drawback is the key management complexity. The risk of key leakage or compromise. which could result in the exposure of sensitive model parameters undermines the goals of privacy-protection.

- **Exploring Homomorphic Encryption and Differential Privacy Techniques towards Secure Federated Learning Paradigm**
  - Authors-Rezak Aziz, Soumya Banerjee, Samia Bouzefrane and Thinh Le Vinh
  - Year Published - 2023
  - Drawbacks that are possible are balancing between differential privacy and Homomorphic encryption can be computationally expensive. This paper also gave an idea on future exploration by combinary both Homomorphic encryption and Differential Privacy considering this way has the most significant potential to safeguard data across all participants in federated learning process.

- **Federated Learning for Edge Computing**
  - Authors - Alexander Brecko, Erik Kajati, Jiri Koziorek, Iveta Zolotova
  - Year Published - 2022
  - Proposed idea - This paper provides an overview of the methods used in Federated Learning with a focus on edge devices with limited computational resources. This paper also presents Federated Learning frameworks that are currently popular and that provide communication between clients and servers.
  - Main drawback is the variance in computational requirements of edge devices such as hardware heterogeneity, communication overload or limited resources of devices.

- **Personalized Federated Learning for Intelligent IoT Applications: A Cloud-Edge Based Framework**
  - Authors - QIONG WU , KAIWEN HE, AND XU CHEN
  - Year Published - 2020
  - Proposed idea - Proposed PerFit, a personalized federated learning framework in a cloud-edge architecture for intelligent IoT applications with data privacy protection. PerFit enables to learn a globally-shared model by aggregating local updates from distributed IoT devices and leveraging the merits of edge computing.
  - Main challenges of federated learning in iot environments are device heterogeneity, caused by varying hardware and network conditions, leads to communication complexities and performance issues. Statistical heterogeneity arises from non-IID data distributions, impacting feature and label consistency.


Cloud
Global model is learned by aggregating and averaging local models
Offloading stage
Learning stage
Personalization stage
Local model
Global model
Global model
Local model
Edge
Integrated communication and computing
Computing tasks
Computing tasks
Global model
Personalized model
Personal information
IoT devices

- **Adaptive Federated Learning in Resource Constrained Edge Computing Systems**
  - Authors -Shiqiang Wang , Tiffany Tuor, Theodoros Salonidis, Kin K. Leung, Christian Makaya, Ting He and Kevin Chan
  - Year Published - 2019
  - Proposed idea - Proposed a control method to achieve the desirable trade-off between local update and global aggregation in order to minimize the loss function under a resource budget constraint.
  - Main drawback is Data imbalance i.e The data distribution across edge nodes may not be uniform, leading to biased models and Edge dynamics i.e, The edge nodes may have different computing power, storage capacity, and communication bandwidth, which can affect the performance of the learning technique .

- **Personalized Federated Learning with Server-Side Information**
  - Authors- JAEHUN SONG, MIN-HWAN OH AND HYUNG-SIN KIM
  - Year Published - 2022
  - Proposed idea - The paper introduces a novel approach, FedSIM, in the context of personalized federated learning with server data. FedSIM optimizes model performance and reduces client computational overhead by actively involving the server in the training process. It accomplishes this by employing meta-gradients calculated on the server,optimization with a custom loss function, followed by higher order gradient computations using server data.
  - Drawbacks are it introduces server dependency, raising concerns about single points of failure and data privacy. The method may demand substantial server resources, adding complexity and scalability challenges.

- **Adaptive privacy-preserving federated learning**
  - Authors - Xiaoyuan Liu, Hongwei Li, Guowen Xu, Rongxing Lu, Miao He
  - Year Published - 2019
  - Proposed idea - This paper aims to address security and privacy concerns in federated deep learning, by improving the balance between privacy and accuracy and develop a framework that applies techniques such as noise injection and relevance assessment, to maintain a strong level of data privacy while minimizing the impact on the accuracy of the deep learning model.
  - Main drawback is the complexity and overhead caused likely due to the introduction of privacy-preserving techniques, which can impact the computational and communication requirements of the federated deep learning process.

Existing System

1. **Traditional Learning Environments:** Many educational institutions still rely heavily on traditional classroom-based learning. While this approach has its merits, it often lacks adaptability to individual student needs. The one-size-fits-all model may not cater to diverse learning styles and paces.
2. **LMS (Learning Management Systems):** Learning Management Systems are widely used in both traditional and online education. These platforms offer content delivery, assessments, and tracking tools. However, they can be rigid in terms of personalization, and the user interfaces may not always be intuitive.

3. **Online Courses and MOOCs:** Massive Open Online Courses (MOOCs) have gained popularity for providing accessible education. However, they can lack personalized interaction and often have low completion rates due to a lack of engagement
4. **Limited Adaptability:** The lack of personalization means that the system doesn't adapt to the specific needs and abilities of each student. Students who need more time to grasp a concept or those who are ready to advance may not get the support they require.

Problem Statement

“Existing educational technology platforms lack privacy protection, hindering their ability to provide adaptive, personalized learning. Our project introduces a privacy-protected Federated Adaptive Learning Platform that leverages encryption and federated learning to redefine education technology, offering tailored, secure learning experiences for students.”

Input: Our EdTech Platform project receives various educational data sources, including student profiles, learning materials, assessments, performance metrics, edge device data, local user data, and model updates.

Output: The primary output is personalized educational recommendations, including learning materials, quizzes, assessments, pathways, progress tracking, and insights for students.

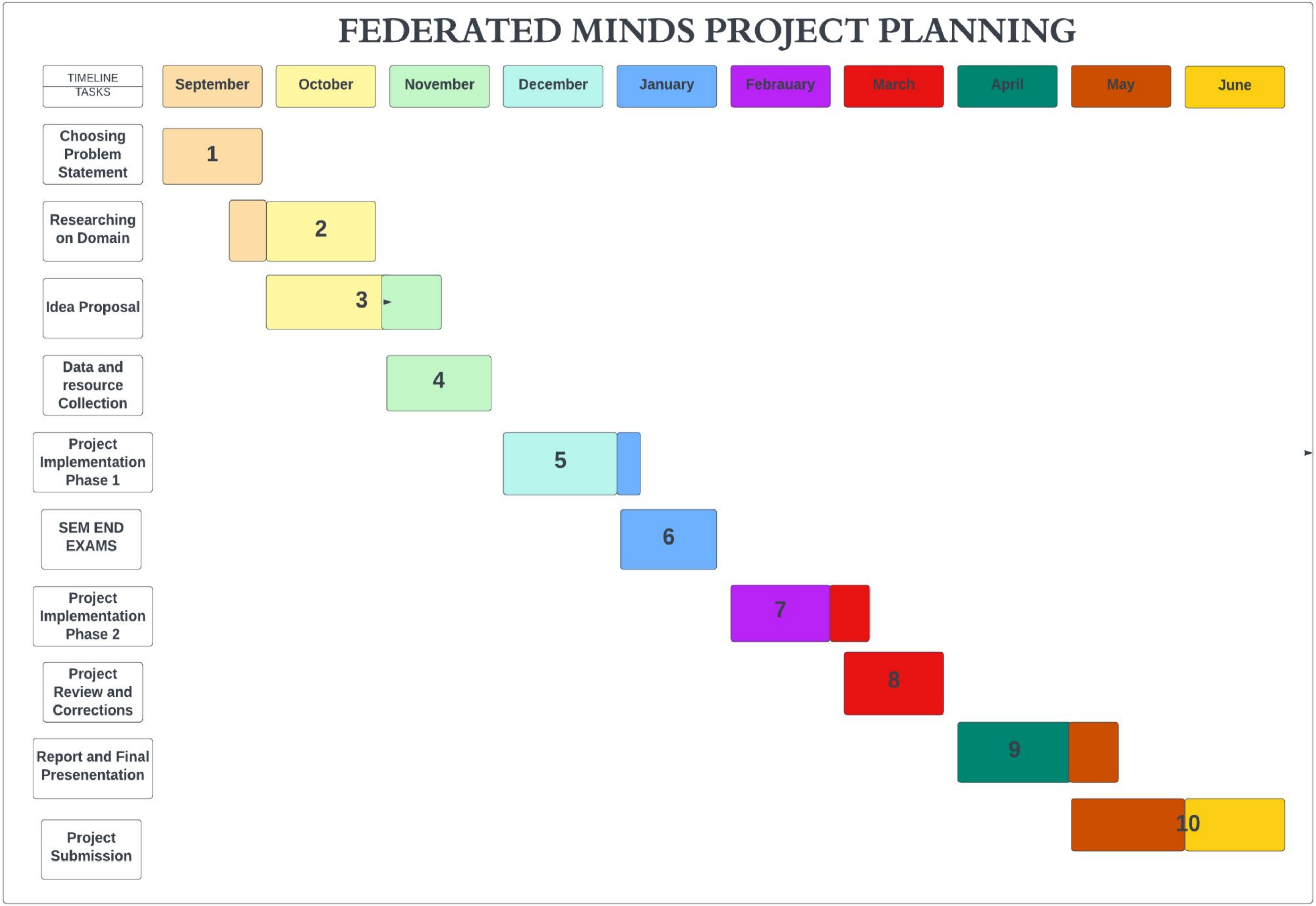
Proposed System

Our "Privacy Protected Federated Adaptive Learning Platform"leverages federated learning, advanced encryptions like homomorphic encryption, to provide personalized, secure education. It adapts to individual performance, offers a user-friendly interface, and scales to meet diverse educational needs.

- **Front End**: Our platform's user interface is built using React, providing a responsive and user-friendly experience for educators, students, and administrators in all browser supporting devices.

- **API** : We employ socket-based communication for real-time interactions and JSON for data exchange, ensuring efficient and interactive user experiences.
- **Encryption**: The system utilizes homomorphic encryption techniques to secure data at rest and in transit, allowing computations on encrypted data without compromising its privacy.
- **Backend Architecture**: Our backend leverages Flask, OpenFL and TensorFlow Federated frameworks in Python, for implementing federated learning on the data collected from the front end and model stacking using ensembling methods on the model vectors obtained from front end. This ensures efficient processing of user data and model updates while maintaining data privacy.

Project Planning/Scheduling



FEDERATED MINDS PROJECT PLANNING

TIMELINE

TASKS

September

October

November

December

January

Febrauary

March

April

May

June

Choosing Problem Statement

1

Researching on Domain

2

Idea Proposal

3

Data and resource Collection

4

Project Implementation Phase 1

5

SEM END EXAMS

6

Project Implementation Phase 2

7

Project Review and Corrections

8

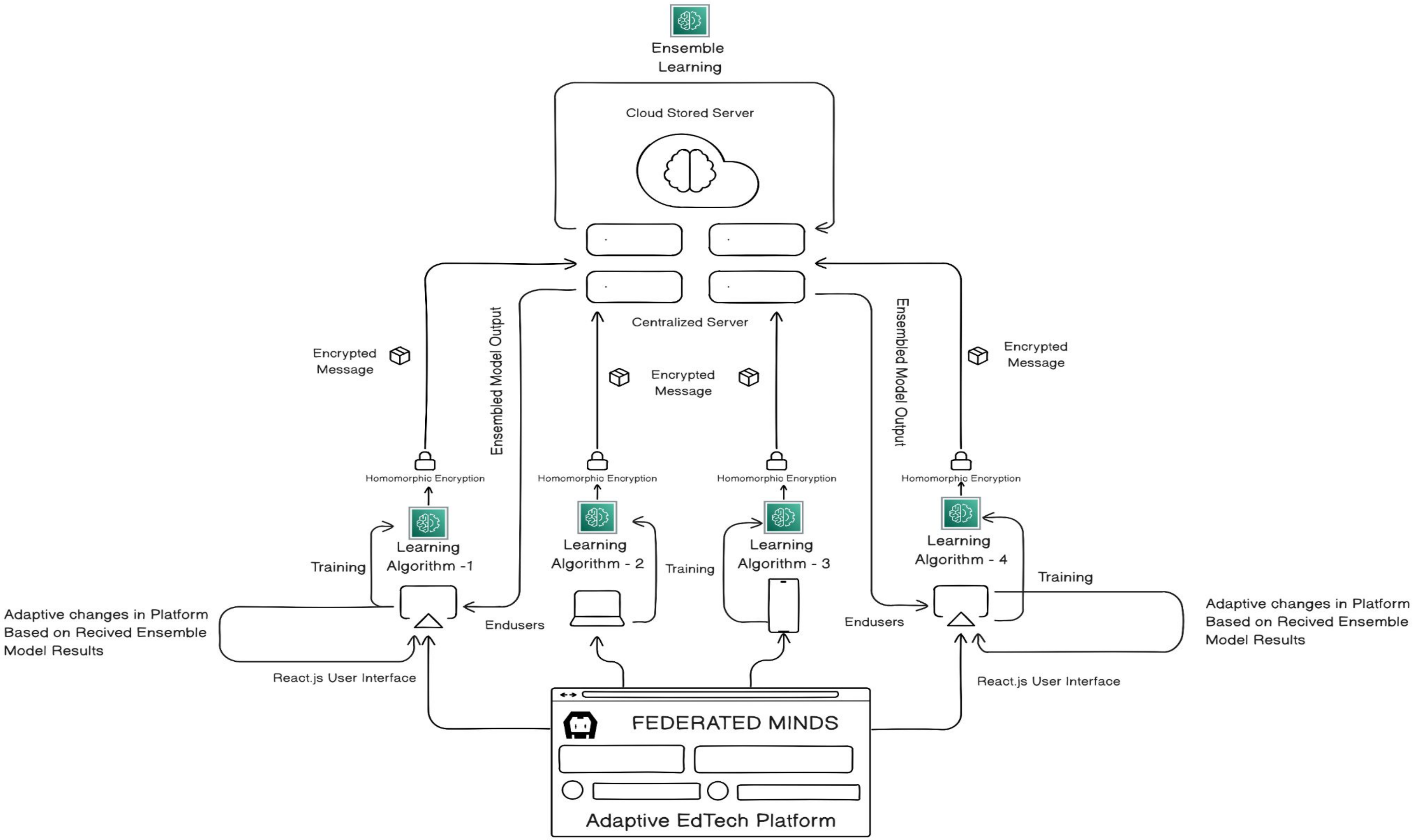
Report and Final Presentation

9

Project Submission

10

Architecture



Applications

1. **Personalized Education**: The platform can be used to create personalized learning experiences for students, adjusting content, pace, and difficulty based on their unique learning styles and abilities.
2. **K-12 and Higher Education Institutions**: Universities and colleges adopt this platform to offer coursework for a diverse student body.
3. **Adaptive Language Learning**: Language learning apps and programs can use the platform to customize lessons based on a learner's progress and proficiency.

4. **Data Security in Educational Institutions**: Schools and universities can ensure the privacy of student data, reducing the risk of data breaches.
5. **Global Access to Education**: The platform can facilitate remote learning for students worldwide, enabling access to education with limited resources.
6. **EdTech Startups and Product Development**: startups can use the platform to create innovative, data-secure educational products and services.

References

[1]Kurniawan, H.; Mambo, M. Homomorphic Encryption-Based Federated Privacy Preservation for Deep Learning. Entropy **2022**, *24*, 1545. https://doi.org/10.3390/e24111545

[2] Xia, Q., Ye, W., Tao, Z., Wu, J. and Li, Q., 2021. A survey of federated learning for edge computing: Research problems and solutions. High-Confidence Computing, 1(1), p.100008.

[3]Kurniawan, H. and Mambo, M., 2022. Homomorphic Encryption-Based Federated Privacy Preservation for Deep Active Learning. Entropy, 24(11), p.1545.

[4] Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R. and Zhou, Y., 2019, November. A hybrid approach to privacy-preserving federated learning. In Proceedings of the 12th ACM on artificial intelligence and security (pp. 1-11).

[5] Wu, Q., He, K. and Chen, X., 2020. Personalized federated learning for intelligent IoT applications: A cloud-edge based framework. IEEE Open Journal of the Computer Society, 1, pp.35-44.

[6] T. Li, A. K. Sahu, A. Talwalkar and V. Smith, "Federated Learning: Challenges, Methods, and Future Directions," in IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50-60, May 2020.

[7] Brecko, A., Kajati, E., Koziorek, J. and Zolotova, I., 2022. Federated learning for edge computing: A survey. Applied Sciences, 12(18), p.9124.

[8] Muñoz, A., Ríos, R., Román, R. and López, J., 2023. A survey on the (in) security of trusted execution environments. Computers & Security, 129, p.103180.

[9] Liu, X., Li, H., Xu, G., Lu, R. and He, M., 2020. Adaptive privacy-preserving federated learning. Peer applications.