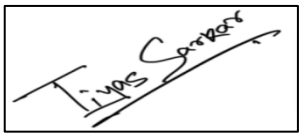
Annexure3b- Complete filing

INVENTION DISCLOSURE FORM

Details of Invention for better understanding: Cloud-Based AI Model Deployment System – A scalable platform for deploying and managing AI models.

1. TITLE: Self-Sufficient Charging System for Electric Public Transportation

2. INTERNAL INVENTOR(S)/ STUDENT(S): All fields in this column are mandatory to be filled

A. Full name	**Tiyas Sarkar(STUDENT 1)**
Mobile Number	6294978581
Email (personal)	info.tiyasofficial11901657@gmail.com
UID/Registration number	12312942
Address of Internal Inventors	School Of Computer Science and Engineering Lovely Professional University, Punjab-144411, India
Signature (Mandatory)	
B. Full name	(FACULTY NAME)
Mobile Number	
Email (personal)	
UID/Registration number	
Address of Internal Inventors	School Of Computer Science and Engineering Lovely Professional University, Punjab-144411, India
Signature (Mandatory)	

C. Full name	**(STUDENT 2)**
Mobile Number	
Email (personal)	
UID/Registration number	
Address of Internal Inventors	School Of Computer Science and Engineering Lovely Professional University, Punjab-144411, India
Signature (Mandatory)	
D. Full name	**Sharanya Vemula(STUDENT 3)**
Mobile Number	7989322217
Email (personal)	sharanyavemula13@gmail.com
UID/Registration number	12218300
Address of Internal Inventors	School Of Computer Science and Engineering Lovely Professional University, Punjab-144411, India
Signature (Mandatory)	
E. Full name	**(STUDENT 4)**
Mobile Number	
Email (personal)	
UID/Registration number	

Address of Internal Inventors	School Of Computer Science and Engineering Lovely Professional University, Punjab-144411, India
Signature (Mandatory)	

3. DESCRIPTION OF THE INVENTION:

1. Purpose

The cloud-based AI model is designed to convert the deployment, management and scaling of artificial intelligence (AI) models into a single, cloud-country platform. It responds to the increasing complexity of AI operations with a scalable, automatic solution to deploy and manage the AI model in individual clouds, on-radius and edge environment. Technology is required because businesses depend on machine learning to direct decision making, automation and user experiences.

2. Technical functioning

3. **Cloud-Personal Model Sports**: The system uses a container-based purinogen (eg, doors, kubernets) to package the AI model and all their dependence. This allows the model to be deployed as microservices that can be accessed through REST or GRPC API. These APIs are automatically generated and version-controlled so that they can be easily used by applications and services.
4. **IOT and Edge Support**: The platform can be deployed not only in the cloud but also to edge the equipment (eg, IOT Gateway, Mobile). The system holds a tab on the capacity and performance of each node and provides real -time decisions and estimates in applications with delays obstacles such as smart manufacturing, autonomous vehicles and medical devices.
5. **Intelligent model orchestration**: With orchestation software such as cubernets, the system will automatically score AI services in real time as demand. It is highly available and is a load-balance traffic traffic and can even send requests for the closing point with the nearest node or the lowest delay.
6. **Real Time Monitoring and Analytics**: The built -in telemetry and logging mechanism track invention performance (delay, throwput), error rate, resource consumption and user behavior. The dashboards imagine these matrix in real time, enabling the active issue of model operation and adaptation.
7. **Automatic Retrening Tiger**: System can detect data drift, model performance decline, or user behavioral change, and re-start jobs through human-in-loop workflows. This ensures that the models are accurate and effective over time.
8. **Safe model governance**: Comprehensive role-based access control (RBAC), audit trails and model approval workflows are applied. This ensures that only authorized personnel can only deploy, update or roll the model, and all changes are tracedable.
9. **Multi-cloud and hybrid integration**: The platform is designed to operate in several cloud providers (AWS, Azure, GCP) and is integrated with the on-primesies environment,

making it suitable for hybrid cloud architecture. The model can be transferred or duplicated to clouds for excess, compliance or cost-evidence.

10. Unique Attributes

1. **Scalability**: Designed to score horizontally and vertically, the platform can support hundreds of concurrent models and millions of daily conclusions requests, without performing fall.
2. **Automation**: From packaging to purinogen, monitoring, retraning, system AI automates important stages in the life cycle, which reduces manual efforts and human error significantly.
3. **Multi-friendly architecture**: Supports several teams or commercial units, which deploys models independently with strong isolation, access control and billing/reporting features.
4. **Plug-end-play purpose**: With minimal configuration, users can deploy models in minutes using a simple UI or CLI. The platform abstract infrastructure abstracts complications, making data scientists capable of focusing on modeling.
5. **Continuous improvement**: The system uses telemetry data and machine learning algorithms to continuously improve performance, recommend adaptation, and resume strategies.

11. Conclusion

The cloud-based AI model perpetuation system represents a major jump as to how the AI model is operated on a scale. Cloud-foreign technologies with intelligent automation, multi-cloud compatibility and enterprise-grade regime, platform AI converts the purinogen into a spontaneous, safe and efficient process. This innovation empowers organizations to exploit the entire power of AI to reduce the operational burden and maximize performance and agility.

A. PROBLEM ADDRESSED BY THE INVENTION:

The adoption of Artificial Intelligence (AI) in the enterprise and research environment has accelerated, but deploying, managing and maintaining AI models on a scale shows many important challenges. Cloud-based AI model purinyogen system addresses the following major problems:

1. **Fragmented finance infrastructure**: Organizations often rely on infrastructure ad hoc or manually configured infrastructure to deploy AI models. This leads to a fragmented environment with inconsistent runtime configurations, obstructing scalability, repetition and reliability of AI solutions in teams and platforms.

2. **Standardization and lack of automation**: Most AI models are deployed using custom scripts or manual processes, which show human error, slow down the deployment cycles, and complicate updates and rollbacks. These manual approaches also lead to a lack of version control, making it difficult to track change or reproduce model performance.

3. **Limited scalability and flexibility**: Traditional purinogen pipelines are often tied to a specific cloud or on-arrival infrastructure. As the demand increases, these pipelines automatically struggle on a scale and cannot efficiently manage workloads in various environments such as edge, hybrid clouds, or multi-region setup.

4. **Inadequate surveillance and observation**: Without monitoring integrated logging and performance, organizations have limited visibility in model health, delay and accuracy. This deficiency of observation makes it difficult to detect issues such as data drift, stability of model or unpredictable system behavior.

5. **No retrening or feedback loop**: Once deployed, most AI models work without reaction to the actual world performance. There is often no mechanism based on triggering retrenching, user's behavior, or model falls, which leads to a decline over time.

6. **Safety and governance interval**: The AI model deployed often lacks adequate security controls, such as certification, authority and audit trails. It highlights sensitive systems to risks, and makes it difficult to manage compliance with data rules and internal access policies.

7. **Developer and data scientific obstacles**: Model perineyogen is usually managed by devops or mlops teams, requiring special infrastructure knowledge. This deployment disconnects data scientists from the pipeline, slows the experiment and hinders innovation.

8. Conclusion

The cloud-based AI model sporting system addresses these pressure issues by providing an integrated, automated platform that simplifies deployment, standardize operations, and supports dynamic scaling and monitoring. This enables the continuous integration and distribution of the AI model, improves governance, and ensures that models are always performing better in the production environment. By reducing the gap between model development and production purpose, this system increases the reliability, scalability and access to AI technologies in industries.

B. OBJECTIVE OF THE INVENTION (Provide minimum two)

 Increase operational efficiency: The cloud-based AI model is one of the main objectives of the model the system system to deploy and automate life cycle management. By offering automatic packaging, scaling and monitoring equipment, the platform reduces manual overhead, reduces deployment time, and increases the reliability of AI services in production.

 Adapt resource uses: The objective of the forum is to compute the resources of computational and infrastructure by digging AI workload dynamically based on real -time demand. By wisely

guessing and integrating with multi-cloud and edge resources, it ensures cost-effective use of GPU, CPU and memory in the atmosphere.

 Provide scalability and flexibility facility: A major purpose is to build a solution that is naturally scalable and adaptable. The system supports the deployment in various infrastructure-clouds, hybrids and on-rich-suitable for all sizes of organizations. This can basically adjust the model, users and increasing number of estimates without a decline in performance.

 Developer and data scientific experience improve: The platform is designed to strengthen developers and data scientists by simplifying the individual interfaces through intelligence interfaces and automatic workflows. This enables rapid use, testing and deployment cycles, reducing dependence on Devops teams.

 Ensure governance and safety: Another major objective is to embed the strong governance features within the model perineeration life cycle. Roll-based access control (RBAC) ensures the system secure, obedient and well-written model operations, by applying the version management and audit logs.

 Supporting continuous learning and model optimization: Platforms enable the mechanisms to continuously learn by involving the mechanism to monitor the performance and trigger the workflow based on the flowing or performance decline. This ensures that the deployed models combine with accurate, updated, up-to-date and real-world data patterns.

C. STATE OF THE ART/ RESEARCH GAP/NOVELTY:

Sr. No.	Study	Abstract	Research Gap	Novelty
1	**AI model serving architecture**	This study discovers traditional model services such as tensorflow serving and torture that wrap the model in APIs.	While these devices serve individual models well, they have large-scale multi-cloud deployment, surveillance and lack of integrated support to resume workflows.	The invention offers an end-to-end platform that automatically automatically deployment, scaling and performance tracking in distributed environment.
2	**MLOps pipeline**	This research machine focuses on CI/CD pipelines for learning, emphasizes model	Most pipelines emphasize training and ignore post-finish monitoring, retrenching	The system shuts off this gap by introducing the workflows associated with the use of runtime observability,

Sr. No.	Study	Abstract	Research Gap	Novelty
		training and experimentation.	trigger, or runtime governance.	performance alerts and production by automated retrenching workflows.
3	**AI model management tools**	This equipment provides partial solutions for existing equipment models tracking and version controls such as MLFLOW and DVC.	These devices are not deployed and focused and lack of cloud-country scaling, access control and runtime integration.	This platform expands the version with cloud-purpose automation, real-time rollback and dynamic routing of invention requests.
4	**Edge AI deployment frameworks**	Periny Management Framework Edge AI Discusses the reduction of delays by attacking the data source close to the data source.	Most framework requires complex manual configurations and does not support hybrid sinsitance or centralized monitoring.	The invention supports hybrid cloud-edge purinogen with integrated dashboard and auto-disverse of the edge nodes.
5	**AI Governance and Security**	The AI regime and security recent work examines the model access control and compliance requirements.	Safety measures are often external to the models that serve layers and are rarely integrated with audit trails and use policies.	The invention uses direct policies in RBAC, audit logging, and complete life cycle regime.

Conclusion

The cloud-based AI model sinner ship system fully integrated, cloud-country and automated platforms, by offering the current AI, addresses the major boundaries in purpose and life cycle management research. By uniting deployment, observation, observation, and governance under a system, it eliminates the fragmentation and disabilities of the current equipment and workflow. Its unique combination of scalability, edge/cloud compatibility, and AI-centered surveillance fills significant intervals in existing solutions and empower organizations to deploy AI with greater speed, control and effect.

D. DETAILED DESCRIPTION:

The cloud-based AI model peripinity system is an innovative solution designed to streamline and score the deployment, management and monitoring of artificial intelligence (AI) models in a diverse computing environment. It integrates modern devops principles, cloud-country infrastructure and intelligent automation to reduce operating overheads and maximize model performance and availability.

- **2. System Components**

2.1 Personogen infrastructure

- **Model container**: The model is contained using techniques such as doors, allowing them to be packed with all dependence for frequent execution on various platforms.
- **Orqualing engine**: The system contains container uses orchestration platforms such as Kuberanets to manage lifestyles. It supports high availability for autosking, rolling updates and deployed models.

2.2 Monitoring and Telemetry Stack

- **Matrix Collection**: The system integrates with devices such as Prometheus and Opentalemetry to capture the metrics such as ingestion delay, throwput, resource usage and error rate.
- **Logging**: The centralized logging model enables behavioral and widespread visibility in the system using devices such as Elk Stack (Elasticsearch, Logstash, Kibana).
- **Alerting System**: Real -time alert can be configured on the basis of discrepancies, traffic spikes or model drifts, allowing active maintenance.

2.3 Central Management System

- **Personogen dashboard**: The model offers an integrated interface to manage the management, health matrix view and configure the endpoints and autoskling policies.
- **Model Registry**: A built -in registry maintains a version of each model, metadata, signature and position (eg, staging, production).
- **Access Control and Governance**: Administrators can define the role-based access permissions, apply approval workflows and models can maintain an audit log for life-cycle activities.

- **3. Technical Functionality**

3.1 Wise Pranicogen Process

- **API generation**: When a model is deployed, the system auto-generates and secures the data and use data as well as safe Rest/GRPC API Endpoints.

- **Canary and Blue-Green Personogen Support**: Enables safe rollouts of new model versions, reduces risk during updates and allows comparison of performance between versions.
- **Edge Compatibility**: The model can be pushed on the edge devices with minimal configurations, using a light agent to handle local estimates and periodic syncing.

3.2 Real-Time Data Management

- **Infererance Logging**: Each estimate request and response for future audit and performance reviews can be logged in (with unknown data if necessary).
- **Drift detection**: System concept/data drift detection, trigger information or re -start the workflow monitors input distribution and model output.
- **Performance insight**: Console live live performance dashboard that supports custom KPI, capable of collaborating data scientists and OPS teams to collaborate and improve model behavior.

3.3 Auto-Retaining and Feedback Loop

- **Scheduled Retaining**: Platform can withdraw the model using newly collected data pipelines on time -time, which is triggered by the calendar or trigger by a display drop.
- **Human-in-Loop (HITL) Workflow**: If desired, the retrenching bicycle can be rooted through human review stages for sensitive models, such as those used in health or finance.
- **Continuous integration with ML pipelines**: The system integrates with devices such as github actions, genkins and kubflows for seamless retrenching and reality.
- **4. Unique Features**

4.1 Scalability

- The model supports the horizontal scaling of the model interesting workload based on the platform traffic volume and can accommodate thousands of concurrent models in regions and cloud platforms.

4.2 Enhanced User Experience

- Through simple CLI tools, SDK and web interfaces, developers and data scientists can deploy models without infrastructure expertise, accelerating innovation cycles.

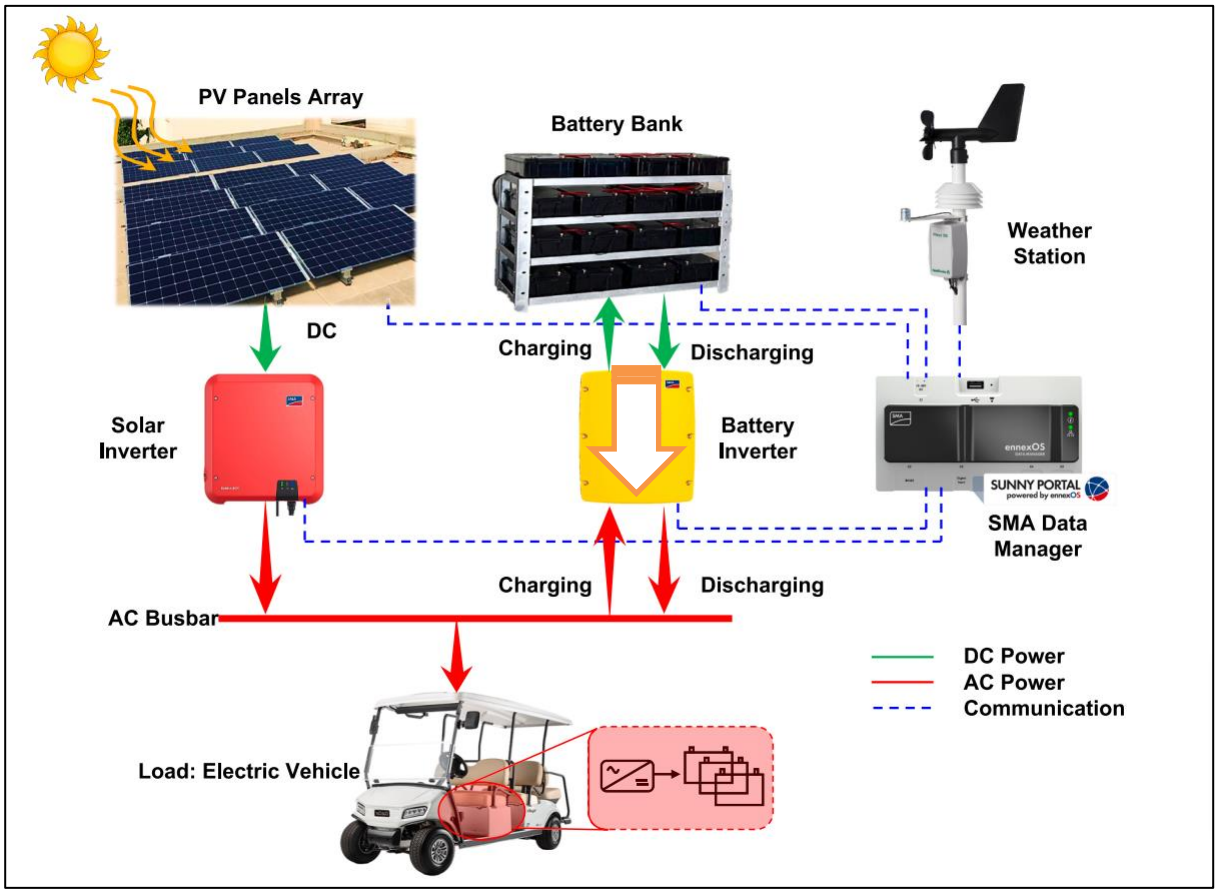
4.3 Low operating complexity

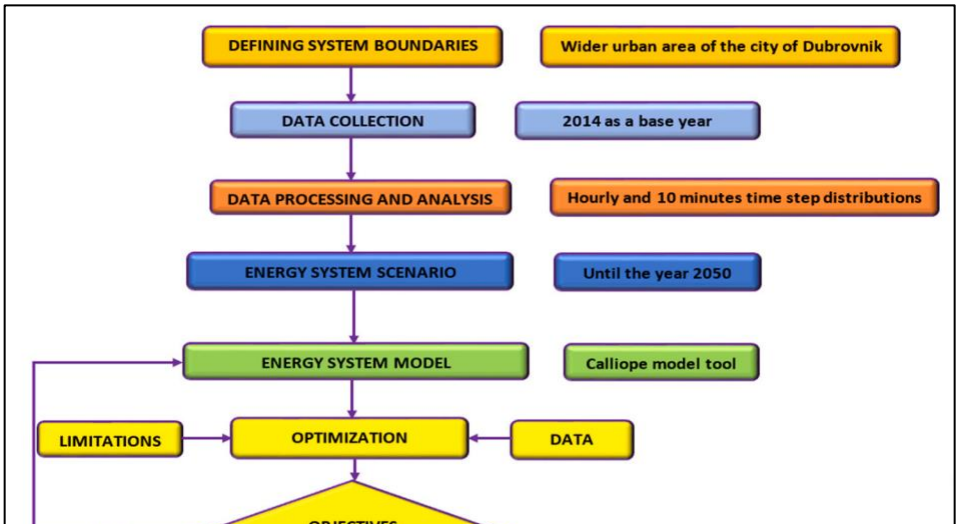
- The built-in automation resources provisions, closing points security, monitoring, and scaling handle-reduces the requirement of dedicated mlops teams.

Conclusion

Cloud-based AI model signal system scalable infrastructure, intelligent monitoring and automated model lifestyle management in an integrated platform makes a significant progress in AI operations by integrating into an integrated platform. Its support for cloud, age and hybrid deployment; Its underlying observation tools; And its flexible retrieving capabilities make it a comprehensive solution to deploy the production-grade AI system. With more security, control and efficiency, reliable production AI enables the system organizations to move from rapid use.

Process Workflow:





E. RESULTS AND ADVANTAGES:

The cloud-based AI model perpetuation system provides many different results and advantages that make it a better solution than the existing AI sinners and life cycle management devices. The main results and benefits of the system below are:

- **1. Increased Operational Efficiency**
- **Low time-to-purpose**: Automatic workflows reduces the time required to deploy and update the AI model, enabled organizations to go from use to production in a matter of minutes.
- **Constant model availability**: Through intelligent load balance and autosaling, the system also ensures high availability and frequent performance of AI services under convertible load conditions.
- **2. Customized resource usage**
- **Dynamic scaling**: The system automatically adjusts resource allocation based on traffic and use, which prevents cloud infrastructure costs from over-provisioning and reducing.
- **Edge-to-cloud flexibility**: The model can be deployed on cloud, hybrid, or edge environment, performing the display and data area based on cases of specific use.
- **3. Better user and developer experience**
- **Real-time monitoring and feedback**: Data scientist and devops team can monitor the model health, delay and use through an intuitive dashboard, which can enable active problems resolving and performing tuning.
- **Simplified operations**: No-code/low-code interface and CLI tools simplify purinioans for users without deep infrastructure knowledge, strengthening data science teams to manage deployment independently.
- **4. Scalability and Flexibility**
- **Multiple support**: Organizations can manage several projects or departments within the same platform with strict separation, user management and access policies.

- **Seamless Integration**: Easily integrate with popular devices such as GIT, Jenkins, Mlflow, Tensorflow, Pytorch and Cloud Providers, making it adapted to diverse technical stacks.
- **5. Maintenance and downtime reduced**
- **Automatic rollback**: In case of humiliated model performance, the system may automatically return to a stable version, reduce service disintegration.
- **Self-healing services**: Integration with orchestration platforms such as Kubernets ensures that model containers are resumed or automatically replaced in terms of failure.
- **6. Increase in security and governance**
- **Roll-based access control (RBAC**): It ensures that only authorized users can only deploy, management or retire the model by supporting enterprise compliance requirements.
- **Audit Trails and Edition Tracking**: Each action - discipline, update, rollback - has been logged on to a model, which provides complete traceability and compliance assistance.
- **Comparison to Existing Prior Art**
- Cloud-based AI model purinyogen system provides many major benefits on traditional purpose methods and available equipment:
- **Ed-Hawk or Scripted Payment**: Manual scripting lacks standardization, observation and error handling. The proposed system automatic these processes with the underlying intelligence and governance.
- **Compared to inheritance models serving devices**: While devices such as tensorflow serving and torturev provide basic API wrappers, they do not resume scalable orchestration, monitoring, or integrated capabilities in the same platform.
- **Classic CI/CD pipelines**: Traditional devops pipelines have not been designed for specific requirements such as data drift detection, model version
- **Conclusion**

The cloud-based AI model is a leading solution for modern AI operations as a leading solution for modern AI operations. This crosses the important boundaries of current systems by providing cloud-country, scalable and intelligent platform for the AI model and management. Its attention on automation, real -time monitoring, model regime and continuous improvement makes an indispensable tool for any organization that demands to operate AI effectively and continuously.

E. EXPANSION:

To ensure comprehensive adoption and effective implementation of the cloud-based AI model needs, several major variables must be considered. These factors affect the design, scalability, integration and overall success of the system in various industries and operating contexts:

- **1. Model and framework compatibility**
- **Model format and framework**: The platform should support a wide range of AI/ML framework such as Tensorflow, Pitorch, ONNX, Skicit-Lurn and XGBost. Compatibility ensures spontaneous onboarding for various AI teams and applications.
- **Serialization standard**: Standardized model formats (eg, .PB, .PT, .PKL, .Onx) help to ensure portability and fertility in the atmosphere.
- **2. Periphery infrastructure**
- **Cloud and Edge Integration**: The system should support deployment in public cloud platforms (eg, AWS, Azure, GCP), private clouds and age devices. This flexibility is essential for delayed-sensitive applications such as IOT, retail or industrial automation.
- **Resources Provisions and Scaling**: Educated resource orchestration (eg, through Kuberanets) ensures optimal use of CPU/GPU resources in all deployment goals.
- **3. Monitoring and communication structure**
- **Observancy tool**: Integration with standardized monitoring tools (Prometheus, Grafana, Alk Stack) is necessary to collect and imagine real -time data.
- **Communication Protocol**: REST, GRPC, Websocket, and MQTT protocols should be supported for versatile integration with client apps, web services and microorvis architecture.
- **4. Data Analytics and AI Operations**
- **Performance tracking**: The system should track accuracy, delay and resource consumption, allowing proper performance tuning and adaptation.
- **Drift detection and retrieving**: The underlying analytics should support the model drift and performance decline detection, and rating the workflow or alert when the threshold is violated.
- **5. Entrepreneurial integration**
- **CI/CD Integration**: Devops is important to automate seamless compatibility with pipelines (eg, Jenkins, Github Activities, Gitlab CI) tests, sins and updates.

- **Identification and Access Management (IAM)**: Integration with enterprise SSO solutions (eg, okta, azure ad) enables secure user authentication and role-based access control.
- **6. Developer and user experience**
- **Self-Seva UI and CLI Equipment**: A user-friendly interface and CLI equipment help scientists and engineers to deploy, monitor and manage models without the need for platform-level expertise.
- **SDK and API**: Java support integration with python, platform-specific SDK, and client application for JavaScript enables developers to create custom interfaces and automation scripts.
- **7. Support and maintenance**
- **Service-level agreement (Slas)**: Production-grade enterprise uptime, delay and clearly defined slas around the response time.
- **Training and documentation**: Onboarding materials, API documentation and customer aid offers ensure smooth adoption and long -term use.
- **8. Regulatory and security compliance Data privacy and governance**: The system should follow rules like GDPR, HIPAA and ISO 27001 to ensure safe data handling and auditability.
- **Model clarity and impartiality**: Equipment to the tool model behavior to audit, the organizations meet moral AI standards and transparency goals.
- **Conclusion**

By accounting for these important variables, the cloud-based AI model is implemented effectively in various industries, operating parameters and technical environment. Its modular and standard-operated architecture allows for widespread compatibility, reliable performance and user-centric scalability. Keeping these factors in mind, it will promote adoption, operational excellence and innovation, able to easily infection in deployment of real world effectively by AI using AI.

F. WORKING PROTOTYPE/ FORMULATION/ DESIGN/COMPOSITION:

Working prototype is not ready. It will take at least a year to complete it.

G. EXISTING DATA:

For the initial prototypes and verification of the cloud-based AI model perpetuation system, a publicly available sample dataset and model repository will be used from platforms such as Hugging Face, Tensorflow Hub and Kagal. The required benchmarking and documentation will be done during the pilot phase.

To effectively support the proposed system, it is necessary to take advantage of existing research and industry data that highlights the operational challenges in AI model siblings, benefits of scalable cloud-based solutions and performance limitations of current devices. Below are the major data categories and supporting evidence.

- **1. Display of AI Paree's Systems**
- **Flammies and Throoput Matrix**: Research by Google (via Tensorflow Serving) and Meta (via torture) suggests that contained AI Infererance services can handle thousands of requests per second with latitude when the AI Infererance services are properly adapted.
- **Case Study-Malty-Model Servator**: Triton Invention Server Benchmark of Nvidia showed performance benefits of up to 30% while serving several models simultaneously using GPU sharing.
- **2. Scalability and infrastructure utilization**
- **Cloud Autoscaling Benefits**: Published by Microsoft Azure and AWS suggests that autockeling resource reduces wastage by 40%based on real -time traffic, making AI model a strong case for dynamic orchestration in serving.
- **Multi-cloud efficiency**: Gartner reports that enterprises using multi-cloud strategies reduce lock-in risk and improve the availability of service up to 35%, which highlight the importance of platform-unknown deployment support.
- **3. Cost and operational efficiency**
- **Total cost of ownership (TCO)**: According to the MCKINSEY AI index, companies using standardized AI deployment platforms reduce their TCO by 20-30% for low maintenance, rapid recurrence and low infrast structure overhead.
- **Model Downtime Impact**: Research from Institute for Operations Research and Management Sciences (Infors) indicates that brief model outs in fields such as Finance or Health Services can also lead to significant service decline and financial loss - reliable, lowering the need for platforms serving platforms.
- **4. Developer and user experience data**
- **Mlops survey data**: In 2023 stack overflow developer survey, 48% AI physicians cited the complexity of deployment and cited tooling deficiency as major obstacles for the

operation of machine learning models. It highlights the need for a proposed such as low-codes, self-service finance systems.

- **Time-to-Parineeration Benchmark**: Github and Google AI's data suggests that traditional model percency workflows take an average of 3-6 weeks on an average from development to production. With automated perineeogen pipelines, it can be less than 1 hour-reduce the operating speed of the cloud-country platforms.
- **5. Comparative analysis of existing technologies**
- **Monolithic vs. modular platform**: A study published in the journal of cloud computing suggests that the moduular AI systems (eg, contained microsarvis) provide better mistake isolation, rapid upliftment and easy scaling than monolithic AI framework.
- **Monitoring Integration**: A review by the Mlops community suggests that AI deployment less than 30% in production is actively monitored for accuracy or drift. Add the integrated monitoring of the proposed platform and directly to the pipelines.
- **Conclusion**

Educational studies, industry reports and benchmark evaluation provide strong support for cloud-based AI model perpetuation system. The advantages of the system in speed, costs, scalability and observation are reflected in the current limitations of all existing equipment and pressure requirements of modern AI-managed organizations. The inclusion of these insights into the development and verification of the system will further strengthen its real -world value and competitive edge in the mlops ecosystem.

4. USE AND DISCLOSURE (IMPORTANT): Please answer the following questions:

A. Have you described or shown your invention/ design to anyone or in any conference?		NO (No)
B. Have you made any attempts to commercialize your invention (for example, have you approached any companies about purchasing or manufacturing your invention)?		NO (No)
C. Has your invention been described in any printed publication, or any other form of media, such as the Internet?		NO (No)
D. Do you have any collaboration with any other institute or organization on the same? Provide name and other details.		NO (No)

E. Name of Regulatory body or any other approvals if required.		NO (No)

5. Provide links and dates for such actions if the information has been made public (Google, research papers, YouTube videos, etc.) before sharing with us. **NA**

6. Provide the terms and conditions of the MOU also if the work is done in collaboration within or outside university (Any Industry, other Universities, or any other entity). **NA**

7. Potential Chances of Commercialization. **Yes**

8. List of companies which can be contacted for commercialization along with the website link.

There are two companies here which specializes in machine learning operations (MLOps) and scalable AI infrastructure. These organizations can be a potential partner for the commercialization of the cloud-based AI model percene system:

1. **Weights & Biases(WandB)**
    - **Overview**: Weight and Biases is a major MLOps company that provides equipment for tracking, model version and collaborative workflow. Their interest in increasing deployment capabilities is closely aligned with the objectives of this invention.
    - **Website**: Wandb
2. **ClearML**
    - **Overview**: ClearML model provides an end-to-end MLOps platform centered on model orchestration, purification and life cycle management. Their infrastructure may be complemented or integrated with this invention for rapid scaling and commercialization.
    - **Website**: CLEAR

9. Any basic patent which has been used and we need to pay royalty to them.

10**. FILING OPTIONS:** Please indicate the level of your work which can be considered for provisional/ complete/ PCT filings - (Provisional)

11. **KEYWORDS:**

- **Cloud-Based AI Deployment**
- **Scalable AI Infrastructure**
- **Model Serving Platform**
- **MLOps**
- **AI Model Lifecycle Management**
- **Inference API**
- **Model Versioning**
- **Auto-Retraining**
- **CI/CD for AI**
- **Edge Deployment**
- **Real-Time Monitoring**
- **AI Governance**
- **Secure AI Operations**
- **Multi-Cloud AI Platform**
- **Data Drift Detection**
- **Model Performance Analytics**
- **Self-Healing AI Services**
- **Continuous AI Integration**
- **AI Deployment Automation**
- **Containerized Model Deployment**

(Letter Head of the external organization)

NO OBJECTION CERTIFICATE

This is to certify that University/Organization Name or its associates shall have no objection if Lovely Professional University files an IPR (Patent/Copyright/Design/any other…….) entitled "……." including the name(s) of,.....…as inventors who is(are) student(s)/employee(s) studying/ working in our University/ organization.

Further Name of the University/Organization shall not provide any financial assistance in respect of said IPR nor shall raise any objection later with respect to filing or commercialization of the said IPR or otherwise claim any right to the patent/invention at any stage.

(Authorised Signatory)