UMHackathon 2025 Judging Criteria
Domain 2: Financial Technology (by **Balaena Quant**)

Preliminary Round

Product: The actual minimum viable product **or** architecture diagram of the framework/developer library tools for machine learning backtesting.

1. Innovation and Originality (20%)

- **Poor (0-4 points)**: The model lacks novelty and is heavily based on existing approaches without meaningful improvements.
    - **0.0-1.0 points**: The model is a direct implementation of an existing algorithm without any modification or adaptation to the problem at hand.
    - **1.1-2.0 points**: The model applies minor changes to a well-known approach but does not introduce meaningful innovations.
    - **2.1-3.0 points**: Some modifications are made, but the methodology remains largely derivative and does not significantly improve upon standard solutions.
    - **3.1-4.0 points**: The model shows minor originality in feature selection, preprocessing, or tuning but does not present new ideas in model architecture or methodology.
- **Fair (4-8 points)**: The model incorporates some innovative aspects but still relies heavily on conventional techniques.
    - **4.1-5.0 points**: The model introduces some novel feature engineering, preprocessing, or parameter tuning, but its core methodology remains standard.
    - **5.1-6.0 points**: The model adapts an existing approach in a meaningful way, such as applying HMM (Hidden Markov Models), CNN (Convolutional Neural Networks), or NLP (Natural Language Processing) in a non-traditional setting, but lacks a fundamentally new perspective.
    - **6.1-7.0 points**: The model combines multiple existing techniques in a novel way, but its innovation is incremental rather than transformative.
    - **7.1-8.0 points**: The model demonstrates originality in certain aspects, such as loss function design, hybrid architectures, or domain adaptation, but does not represent a major conceptual shift.
- **Good (8-12 points)**: The model presents a creative and original approach that enhances or extends existing methodologies.
    - **8.1-9.0 points**: The model significantly improves upon existing techniques, such as introducing a novel training strategy or optimising an HMM for previously unexplored applications.

- ○ **9.1-10.0 points**: The model applies CNNs or NLP techniques in a unique manner, such as custom feature extraction for time-series data or innovative sequence modelling.
  - ○ **10.1-11.0 points**: The model introduces a novel framework or combination of ML techniques that demonstrate clear advantages over traditional methods.
  - ○ **11.1-12.0 points**: The model successfully applies an unconventional approach, such as merging probabilistic models with deep learning, to achieve superior performance.
- **Very Good (12-16 points)**: The model demonstrates strong innovation and presents a fresh approach that is significantly different from conventional methods.
  - ○ **12.1-13.0 points**: The model proposes a novel adaptation of HMM, CNN, or NLP that effectively improves performance beyond existing benchmarks.
  - ○ **13.1-14.0 points**: The model introduces a unique combination of methodologies, such as an interpretable hybrid model that bridges probabilistic inference with deep learning.
  - ○ **14.1-15.0 points**: The model introduces fundamentally new components, such as an innovative training paradigm, self-supervised learning mechanism, or an improved state representation in HMMs.
  - ○ **15.1-16.0 points**: The model significantly advances the state-of-the-art by proposing novel architectures, loss functions, or optimisation techniques that outperform conventional methods.
- **Excellent (16-20 points)**: The model is groundbreaking, offering a novel contribution that redefines the approach to the problem.
  - ○ **16.1-17.0 points**: The model fundamentally changes how the problem is addressed, introducing a novel machine learning framework or an innovative way to leverage existing architectures.
  - ○ **17.1-18.0 points**: The model introduces an entirely new algorithmic perspective, such as redefining HMM transitions with deep learning-based state representations or introducing new feature extraction techniques for NLP/CNN.
  - ○ **18.1-19.0 points**: The model represents a significant breakthrough, introducing a completely novel ML methodology that challenges existing paradigms.
  - ○ **19.1-20.0 points**: The model is transformative, setting new benchmarks and potentially defining a new standard in the field.

2. Technical Execution and Feasibility (20%)

- **Poor (0-4 points)**: Prototype does not meet basic functionality requirements or fails to demonstrate key features.
  - **0.0-1.0 points**: The prototype does not meet basic functionality and fails to demonstrate key features.
  - **1.1-2.0 points**: The prototype lacks essential features and does not work as intended.
  - **2.1-3.0 points**: The prototype works with limited functionality but fails to meet critical requirements.
  - **3.1-4.0 points**: The prototype demonstrates some core features but lacks key functionalities or does not work correctly.
- **Fair (4-8 points)**: The prototype works but with some issues, and may lack certain key functionalities.
  - **4.1-5.0 points**: The prototype works somewhat but is missing key functionalities or has significant bugs.
  - **5.1-6.0 points**: The prototype is functional but lacks some key features or has major usability issues.
  - **6.1-7.0 points**: The prototype is mostly functional, demonstrating core features, but still requires substantial improvement.
  - **7.1-8.0 points**: The prototype is functional but lacks certain key features and needs significant improvement for full usability.
- **Good (8-12 points)**: The prototype functions well and demonstrates the core features, but it may not be fully optimised or refined.
  - **8.1-9.0 points**: The prototype functions well and demonstrates most core features, though some refinements are needed.
  - **9.1-10.0 points**: The prototype demonstrates core features effectively, though there is room for optimisation or refinement.
  - **10.1-11.0 points**: The prototype is mostly refined, demonstrating core functionality but may still require minor adjustments.
  - **11.1-12.0 points**: The prototype works well, demonstrating core features with minimal issues, though it may still require optimisation for performance.
- **Very Good (12-16 points)**: The prototype is robust, efficiently demonstrating the core functionality and addressing the problem effectively.
  - **12.1-13.0 points**: The prototype is robust, demonstrating core functionality efficiently with minor issues.
  - **13.1-14.0 points**: The prototype is efficient, demonstrating core functionality and addressing the problem effectively with optimised performance.
  - **14.1-15.0 points**: The prototype is well-executed, demonstrating the core features with high efficiency and addressing the problem effectively.
  - **15.1-16.0 points**: The prototype demonstrates robust functionality, addressing the problem effectively with an optimised and polished design.
- **Excellent (16-20 points)**: The prototype is well-executed, polished, and operates flawlessly, with an optimised design that is scalable and practical.

- ○ **16.1-17.0 points**: The prototype is well-executed with an optimised design, operating flawlessly and scalable for practical use.
- ○ **17.1-18.0 points**: The prototype is polished, demonstrating flawless operation with optimised design, and ready for real-world deployment.
- ○ **18.1-19.0 points**: The prototype is exceptionally well-executed, demonstrating scalable, practical solutions with flawless operation.
- ○ **19.1-20.0 points**: The prototype operates flawlessly, with an optimised, scalable design ready for large-scale real-world applications.

3. Impact and Usefulness (20%)

- **Poor (0-4 points)**: The solution is irrelevant, not utilised by the team, or fails to measure strategies using the required performance metrics (Sharpe ratio, maximum drawdown, and trades per interval).
    - **0.0-1.0 points**: The solution is entirely irrelevant to the problem statement.
    - **1.1-2.0 points**: The solution has some relevance to the problem statement but lacks practical application.
    - **2.1-3.0 points**: The solution is relevant but is not utilised by the team to generate at least one trading strategy.
    - **3.1-4.0 points**: The solution is utilised by the team to generate strategies, but the strategies are not evaluated using the required performance metrics (Sharpe ratio, maximum drawdown, and trades per interval).
- **Fair (4-8 points)**: The solution generates strategies but lacks completeness, efficiency, or robustness with partially correct backtesting.
    - **4.1-5.0 points**: The solution generates strategies but only measures performance using one of the required metrics. Backtesting might be only partially correct.
    - **5.1-6.0 points**: The solution generates strategies that are measured by at least two required performance metrics. Backtesting might be only partially correct.
    - **6.1-7.0 points**: The solution generates and measures strategies with all required metrics. Backtesting might be only partially correct.
    - **7.1-8.0 points**: The solution generates and measures strategies with required and additional metrics but lacks clarity, consistency, or thorough backtesting.
- **Good (8-12 points)**: The solution generates strategies that are of some practical value with correct backtesting.
    - **8.1–9.0 points**: The solution generates strategies but only measures performance using one of the required metrics, with correct backtesting.
    - **9.0–10.0 points**: The solution generates strategies that are measured by at least two required performance metrics with correct backtesting.
    - **10.1–11.0 points**: The solution generates and measures strategies with all required metrics with correct backtesting.
    - **11.1–12.0 points**: The solution generates and measures strategies with required and additional metrics with thorough backtesting.
- **Very Good (12-16 points)**: The solution generates strategies with less ideal performance with all required strategies and correct backtesting.
    - **12.1–13.0 points**: The solution generates strategies with Sharpe ratio (SR) <= 0 **or** max drawdown (MDD) <= -50% **or** trades per interval < 3%.
    - **13.1-14.0 points**: The solution generates strategies with $0 < SR <= 1$ **and** MDD > -50% **and** trades per interval >= 3%.
    - **14.1-15.0 points**: The solution generates strategies with $1 < SR <= 1.5$ **and** MDD > -50% **and** trades per interval >= 3%.
    - **15.1-16.0 points**: The solution generates strategies with $1.5 < SR <= 1.8$ **and** MDD > -50% **and** trades per interval >= 3%.

- **Excellent (16-20 points)**: The solution generates strategies with ideal performance with all required strategies and correct backtesting.
    - **16.1-17.0 points**: The solution generates strategies with 1.8 < SR <= 2.0 **and** MDD > -50% **and** trades per interval >= 3%.
    - **17.1-18.0 points**: The solution generates strategies with 1.8 < SR <= 2.0 **and** MDD > -30% **and** trades per interval >= 3%.
    - **18.1-19.0 points**: The solution generates strategies with SR > 2.0 **and** MDD > -50% **and** trades per interval >= 3%.
    - **19.1-20.0 points**: The solution generates strategies with SR > 2.0 **and** MDD > -30% **and** trades per interval >= 3%.

4. Presentation and Communication (20%)

- **Poor (0-4 points)**: The presentation is unclear, disorganised, and difficult to follow. The problem and solution are not adequately communicated.
  - **0.0-1.0 points**: The presentation is unclear and disorganised, failing to communicate the problem or solution effectively.
  - **1.1-2.0 points**: The presentation has major clarity issues and disjointed flow, making it hard to follow the problem and solution.
  - **2.1-3.0 points**: The presentation lacks structure, with some key details missing. It's hard to follow or understand the core message.
  - **3.1-4.0 points**: The presentation is confusing and disorganised and fails to convey the key points effectively.
- **Fair (4-8 points)**: The presentation has some clarity but lacks structure and key details, making it somewhat difficult to follow.
  - **4.1-5.0 points**: The presentation has basic clarity but lacks structure or key details, making it difficult to follow.
  - **5.1-6.0 points**: The presentation has some structure, but the clarity and flow need improvement. It is somewhat difficult to follow.
  - **6.1-7.0 points**: The presentation communicates the main points but is still not fully organised. Some sections need more clarity or details.
  - **7.1-8.0 points**: The presentation is clear but lacks full structure or refinement, leading to gaps in understanding.
- **Good (8-12 points)**: The presentation is clear and well-structured, though some areas could be refined for better flow or more details.
  - **8.1-9.0 points**: The presentation is clear and well-structured, though a few areas need refinement or further clarification.
  - **9.1-10.0 points**: The presentation is organised, clear, and flows logically, but some sections could use more details or better conciseness.
  - **10.1-11.0 points**: The presentation is well-structured and clear, with good flow. However, some sections could benefit from greater refinement.
  - **11.1-12.0 points**: The presentation is clear and well-structured, effectively conveying the solution.
- **Very Good (12-16 points)**: The presentation is highly organised, with a clear, concise, and engaging narrative that effectively communicates the problem, solution, and impact.
  - **12.1-13.0 points**: The presentation is highly organised, clear, and communicates the message engagingly.
  - **13.1-14.0 points**: The presentation is well-organised, concise, and engaging, communicating the problem and solution effectively.
  - **14.1-15.0 points**: The presentation is clear, organised, and compelling, engaging the audience while effectively conveying the problem and solution.
  - **15.1-16.0 points**: The presentation is highly engaging, clear, and persuasive, delivering a strong message that resonates with the audience.

- **Excellent (16-20 points)**: The presentation is compelling and persuasive and seamlessly communicates the idea with clarity and engagement, making the solution easy to understand and inspiring confidence.
    - **16.1-17.0 points**: The presentation is dynamic, engaging, and persuasive, clearly conveying the idea with confidence.
    - **17.1-18.0 points**: The presentation is compelling, engaging, and seamless, presenting the solution with clarity and confidence.
    - **18.1-19.0 points**: The presentation is exceptionally clear and persuasive, inspiring confidence and resonating with the audience.
    - **19.1-20.0 points**: The presentation is exceptionally compelling, clear, and highly engaging, leaving a strong, confident impression on the audience.

5. Team Collaboration and Problem-Solving Process (20%)

- **Poor (0-4 points)**: The team demonstrates poor collaboration with significant issues in task division and problem-solving.
    - **0.0-1.0 points**: The team lacks collaboration, with significant issues in task division and solving the problem.
    - **1.1-2.0 points**: The team faces major collaboration issues and poor coordination, resulting in ineffective problem-solving.
    - **2.1-3.0 points**: The team has some collaboration, but there are key problems in task division and problem-solving.
    - **3.1-4.0 points**: The team demonstrates poor collaboration with unclear roles and ineffective problem-solving.
- **Fair (4-8 points)**: The team collaborated moderately well but faced challenges in communication and task division.
    - **4.1-5.0 points**: The team showed moderate collaboration, with communication challenges or unclear task division.
    - **5.1-6.0 points**: The team faced some challenges in task division or communication but worked together to solve the problem.
    - **6.1-7.0 points**: The team demonstrated moderate collaboration, with some issues in task division or coordination.
    - **7.1-8.0 points**: The team displayed adequate collaboration but struggled with coordination and task division.
- **Good (8-12 points)**: The team demonstrated good collaboration, effectively dividing tasks and working together, though minor coordination issues may have occurred.
    - **8.1-9.0 points**: The team demonstrated good collaboration and division of tasks, but minor coordination issues were present.
    - **9.1-10.0 points**: The team worked effectively together, with clear roles and division of tasks, though minor coordination issues occurred.
    - **10.1-11.0 points**: The team displayed good collaboration and clear communication, efficiently solving the problem with minor coordination challenges.
    - **11.1-12.0 points**: The team demonstrated strong collaboration, with well-defined roles and good communication throughout the problem-solving process.
- **Very Good (12-16 points)**: The team worked cohesively with strong communication and efficient problem-solving. Roles were clearly defined, and collaboration was smooth.
    - **12.1-13.0 points**: The team worked cohesively, communicating well and solving the problem efficiently.
    - **13.1-14.0 points**: The team demonstrated strong communication and collaborative efforts, addressing the problem effectively.
    - **14.1-15.0 points**: The team worked seamlessly, with clear roles and strong communication, solving the problem with efficiency.

- ○ **15.1-16.0 points**: The team demonstrated exceptional collaboration with efficient problem-solving, clear roles, and strong communication.
- **Excellent (16-20 points)**: The team demonstrated exceptional collaboration with seamless communication, a well-defined workflow, and efficient problem-solving.
  - ○ **16.1-17.0 points**: The team demonstrated exceptional collaboration with clear roles and seamless communication.
  - ○ **17.1-18.0 points**: The team worked together exceptionally well, solving the problem with efficient communication and task division.
  - ○ **18.1-19.0 points**: The team demonstrated outstanding collaboration, working together seamlessly and solving the problem efficiently.
  - ○ **19.1-20.0 points**: The team worked exceptionally well, demonstrating strong communication, efficient task division, and effective problem-solving throughout.

Final Round

Product: The framework/developer library for machine learning backtesting **and** at least one strategy found using this particular product.

1. Technical Implementation and Architecture (30%)
- **Architecture and Design**
  - **Poor (0-3 points)**: The solution lacks a clear technical architecture, and there is no consideration for scalability or resource optimisation. The design is not aligned with real-world applications.
    - **0.0-1.0 points**: No discernible architecture. The solution is not organised with no clear technical design.
    - **1.1-2.0 points**: Minimal architecture with basic structure. Lacks integration and scalability. The design is overly simplistic.
    - **2.1-3.0 points**: Basic architecture exists but lacks depth. Scalability and performance considerations are not fully explored.
  - **Fair (3-6 points)**: The architecture is defined but lacks depth or scalability. Some components may not be integrated well, or there may be challenges in scaling the solution.
    - **3.1-4.0 points**: The architecture is outlined, but scalability is not well-considered, or the components may not be fully integrated.
    - **4.1-5.0 points**: The architecture is defined and functional, though scalability or optimisation may need more work.
    - **5.1-6.0 points**: Solid architecture with clear consideration for scalability. Some areas need further optimisation for real-world application.
  - **Good (6-9 points)**: The architecture is well-defined, with consideration for scalability and integration. Some improvements are needed for full optimisation.
    - **6.1-7.0 points**: Well-defined architecture, some consideration for scalability, but integration challenges remain.
    - **7.1-8.0 points**: Scalable architecture with well-implemented integration; optimisation could be enhanced for performance.
    - **8.1-9.0 points**: Highly scalable architecture with excellent integration and performance considerations, though minor tweaks are needed.
  - **Very Good (9-12 points)**: The architecture is robust, scalable, and integrates well with modern tools and technologies. There is a clear consideration for performance, scalability, and future growth.
    - **9.1-10.0 points**: The architecture is fully scalable, integrates with existing tools and technologies, and addresses

performance and future growth. Some minor performance refinements could improve its readiness for large-scale deployment.
- **10.1-11.0 points**: The solution is highly optimised and can scale effectively. It demonstrates good integration with core systems and technologies and addresses scalability with high attention to future-proofing.
- **11.1-12.0 points**: Exceptional architecture that is robust, scalable, and efficient. Integration with multiple systems is seamless, and the solution is prepared for handling future challenges and scaling to meet growth demands.
- **Excellent (12-15 points)**: The architecture is highly optimised, scalable, and future-proof. Demonstrates the use of cutting-edge technologies with seamless integration across multiple systems, ensuring both performance and scalability.
  - **12.1-13.0 points**: The architecture is top-tier, integrating seamlessly across various systems with minimal performance issues. It shows clear future-proofing and scalability, built to handle industry-level demands.
  - **13.1-14.0 points**: Exceptional architecture that employs cutting-edge technologies. It demonstrates not only scalability and integration but also robust performance management and high adaptability to future developments.
  - **14.1-15.0 points**: The architecture is industry-leading, built for seamless, large-scale deployment. It integrates flawlessly with multiple systems and demonstrates exceptional performance optimisation. The solution is highly future-proof, optimised for rapid scaling, and prepared for long-term usage in real-world systems.

- **Integration and Real-World Deployment**
  - **Poor (0-3 points)**: The solution is not suitable for integration into existing systems or for real-world deployment. The prototype is incomplete and lacks functionality.
    - **0.0-1.0 points**: The prototype is incomplete or non-functional. No components are integrated, and there is no pathway for real-world deployment.
    - **1.1-2.0 points**: The prototype demonstrates basic functionality, but it is not deployable or integrated with existing systems. Key features are missing, or there are major performance issues.

- 2.1-3.0 points: The prototype is partially deployable but lacks integration with critical systems. Major issues with functionality or performance prevent it from being applicable for real-world deployment.
- **Fair (3-6 points)**: The prototype demonstrates basic functionality but may have integration challenges or is not optimised for real-world use.
    - **3.1-4.0 points**: The prototype is partially functional with basic integration. However, significant challenges exist in ensuring the solution works effectively in real-world environments due to the lack of optimisation and effective integration issues.
    - **4.1-5.0 points**: The solution demonstrates functional features, but its integration with external systems is problematic or incomplete. The solution is not fully optimised for real-world use and may require substantial modifications.
    - **5.1-6.0 points:** The solution is mostly functional and can integrate with some systems. It may still face issues in real-world scenarios as it lacks full scalability or optimisation for large-scale deployment.
- **Good (6-9 points)**: The prototype is deployable with minimal changes and can integrate into existing systems, though additional optimisation may be needed.
    - **6.1-7.0 points**: The prototype is deployable in small-scale real-world scenarios with minimal integration required. Performance and scalability considerations may still require further refinement to handle larger, more complex environments.
    - **7.1-8.0 points**: The solution is functional, with basic integration into real-world systems. Optimisation for performance is still required to address scalability or fine-tune its deployment for higher loads.
    - **8.1-9.0 points**: The solution is well-integrated into existing systems and works well for deployment in real-world environments. Minor tweaks or further optimisation may be needed for large-scale deployment or complex scenarios.
- **Very Good (9-12 points)**: The solution is well-suited for deployment and integration, with clear pathways for scaling and implementing it in real-world scenarios.
    - **9.1-10.0 points**: The solution is fully deployable and integrates well with other systems. The prototype has clear pathways to scale and can handle real-world use effectively with some minor optimisations.
    - **10.1-11.0 points**: The solution is well-optimised for real-world deployment, with clear integration pathways and scalable

architecture. It's ready for early-stage deployment with minimal changes.
- **11.1-12.0 points**: The solution is fully scalable and demonstrates excellent integration with real-world systems. The deployment is ready for a large-scale environment, with performance and scalability clearly addressed.
- **Excellent (12-15 points)**: The solution is fully deployable and integrates seamlessly into existing systems. It is highly optimised for real-world environments and shows significant potential for large-scale adoption.
  - **12.1-13.0 points**: The solution is fully deployable, seamlessly integrating into existing systems. It demonstrates advanced optimisation for large-scale environments and has the capacity to handle significant user traffic or data loads.
  - **13.1-14.0 points**: The solution is industry-ready, optimised for real-world environments, and with perfect integration. It offers high scalability, meets the needs of real-world scenarios, and requires minimal adaptation for large-scale deployment.
  - **14.1-15.0 points**: The solution is exceptional, fully integrated, and optimised for seamless, large-scale deployment. It demonstrates cutting-edge performance, scalability, and integration capabilities, making it ready for immediate adoption in complex, real-world systems.

2. Prototype Quality and Functionality (30%)

- **Machine Learning Modelling and Performance**
  - **Poor (0-3 points)**: The model is ineffective, failing to meet basic validation criteria or exhibiting major flaws in methodology.
    - **0.0-1.0 points**: Validation accuracy is below 0.5, precision and recall are significantly low (<0.3), and the model is unsuitable for use.
    - **1.1-2.0 points**: Validation accuracy is between 0.5 and 0.6, but precision/recall remain below 0.5. The model lacks robustness and generalisability.
    - **2.1-3.0 points**: The model barely meets minimum requirements (accuracy > 0.6, precision/recall ~0.5) but suffers from poor feature selection, lack of proper validation, or flawed assumptions.
  - **Fair (4-6 points)**: The model meets basic performance criteria but lacks refinement in feature selection, hyperparameter tuning, or interpretability.
    - **3.1-4.0 points**: Accuracy exceeds 0.65, but precision/recall remain marginally acceptable (~0.5). Model assumptions or preprocessing steps are not well-documented.
    - **4.1-5.0 points**: Model performance is reasonable (accuracy > 0.7), but feature selection is suboptimal, leading to potential overfitting or instability.
    - **5.1-6.0 points**: The model is functional, but issues such as improper handling of time-series data or lack of justification for hyperparameters persist.
  - **Good (6-9 points)**: The model is well-structured, with proper validation, feature engineering, and risk mitigation for time-series bias.
    - **6.1-7.0 points**: The model meets validation accuracy (>0.7), precision (>0.5), and recall (>0.5) but could improve feature selection or retraining strategy.
    - **7.1-8.0 points**: Time-series constraints are respected (e.g., proper lagging, rolling preprocessing). Feature selection is justified, though some optimisations remain.
    - **8.1-9.0 points**: The model is properly tuned, cross-validated without look-ahead bias, and demonstrates strong generalisability with a logical retraining schedule.
  - **Very Good (9-12 points)**: The model is robust, efficient, and demonstrates strong methodological justifications.

- **9.1-10.0 points**: Assumptions about hidden states (in HMM) or deep learning architecture are explicitly justified, and hyperparameter tuning is well-documented.
- **10.1-11.0 points**: Model performance is optimised (accuracy > 0.75, strong precision/recall balance). Feature selection is well-supported by statistical tests.
- **11.1-12.0 points**: The model is near production-ready, incorporating a retraining strategy, proper time-series validation, and robust documentation of its methodology.
- **Excellent (12-15 points)**: The model is highly optimised, with a well-justified methodology, strong generalisation, and clear real-world applicability.
  - **12.1-13.0 points**: The model is tuned to a high standard, with exceptional accuracy, precision, and recall. It includes robust documentation and clear rationale for all design choices.
  - **13.1-14.0 points**: Model assumptions (HMM transition probabilities, DL (deep learning) hyperparameters) are fully justified, making it interpretable and reliable. Time-series validation is flawless.
  - **14.1-15.0 points**: The model is state-of-the-art for the given problem, demonstrating exceptional performance, reproducibility, and practical viability.

- **Scalability and Real-World Application**
  - **Poor (0-3 points)**: The solution lacks scalability and is not suitable for real-world application. It would not be viable in large-scale environments.
    - **0.0-1.0 points**: The solution is not scalable and not deployable in real-world environments. The design or architecture cannot handle large volumes of users, data, or systems.
    - **1.1-2.0 points**: The solution has some basic functionality but lacks fundamental scalability considerations. It would require a complete redesign to be used in real-world scenarios.
    - **2.1-3.0 points**: The prototype lacks scalability or is incomplete for large-scale deployment. It may have some components working, but major design flaws prevent it from being deployed in a real-world environment.
  - **Fair (3-6 points)**: The solution shows some potential for scalability but is not fully optimised for large-scale deployment.
    - **3.1-4.0 points**: The solution demonstrates basic scalability but is not optimised for real-world environments. It could handle small-scale use, but major challenges exist when scaling to larger systems or user bases.

- **4.1-5.0 points**: The solution works in small environments, but scalability concerns exist for real-world applications. Some parts of the system may not be optimised to handle higher volumes or complex integrations.
- **5.1-6.0 points**: The prototype shows potential scalability but needs significant optimisation and refinement to handle large-scale deployment in complex systems or real-world use cases.
○ **Good (6-9 points)**: The prototype is scalable and could be applied in real-world scenarios with some refinement for larger-scale use.
- **6.1-7.0 points**: The solution is scalable and functional, but some elements require further optimisation for large-scale use. It may work in smaller or controlled environments, but improvements are needed for broader application.
- **7.1-8.0 points**: The prototype demonstrates solid scalability for real-world application. Some minor adjustments or optimisations are needed to support higher user volumes or integrate into complex systems.
- **8.1-9.0 points**: The solution is scalable and works well in real-world environments with minimal adjustments. It demonstrates a clear path toward larger-scale use with some minor refinements needed.
○ **Very Good (10-12 points)**: The solution is highly scalable, and the prototype demonstrates a clear path toward real-world application and widespread deployment.
- **9.1-10.0 points**: The solution is highly scalable, can be applied in real-world scenarios, and is almost ready for widespread deployment. Only minor optimisations are needed to fully integrate into large-scale systems.
- **10.1-11.0 points**: The prototype is well-suited for large-scale deployment with clear strategies for scaling and real-world application. Optimisation may still be necessary, but the path forward is clear and practical.
- **11.1-12.0 points**: The solution is ready for deployment in large-scale systems. It demonstrates high scalability and shows how it can be implemented effectively in real-world applications across various industries.
○ **Excellent (12-15 points)**: The solution is built with scalability and real-world applications in mind. It demonstrates potential for deployment at scale and integration into large systems, with robust support for growth.
- **12.1-13.0 points**: The solution is industry-ready, with scalable architecture and real-world application. It integrates seamlessly with existing systems and has robust support for future growth and large-scale deployment.

- **13.1-14.0 points**: The solution is highly optimised for both scalability and real-world deployment. It handles large-scale traffic and integrates into complex systems. The solution is ready for immediate adoption in the industry.
- **14.1-15.0 points**: The solution is exceptional, demonstrating seamless integration and unmatched scalability. It is optimised for large-scale, real-world application and is future-proofed for rapid growth, making it ready for global deployment.

3. Code Quality, Documentation, and Maintainability (20%)

- **Code Quality and Modularity**
  - **Poor (0-2 points)**: The code is unstructured, inefficient, and difficult to maintain. There is minimal modularity, making future updates or changes challenging.
    - **0.0-0.5 points**: The code is completely unstructured, difficult to follow, and lacks basic modularity. Significant refactoring is needed, and future updates would be very challenging.
    - **0.6-1.0 points**: The code is inefficient and lacks modularity. It's hard to maintain, and critical design flaws would require major restructuring for scalability.
    - **1.1-1.5 points**: The code has some basic structure but is inefficient and lacks modularity. Future changes or scaling would require significant work.
    - **1.6-2.0 points**: The code is functional, but it is not optimised and lacks sufficient modularity for easy maintenance. The overall design needs major improvements for future growth.
  - **Fair (2-4 points)**: The code is somewhat modular but contains inefficiencies or redundancies. It is difficult to maintain without significant refactoring.
    - **2.1-2.5 points**: The code has basic modularity, but several parts lack optimisation and include redundant sections, making it difficult to maintain.
    - **2.6-3.0 points**: The solution has some modular components but lacks efficiency or performance optimisation. Minor inefficiencies make the code harder to scale.
    - **3.1-3.5 points**: The code is somewhat structured but still needs substantial refactoring due to redundancies or performance bottlenecks.
    - **3.6-4.0 points**: The code is functional, but there are modular issues, inefficiencies, or organisational flaws that need to be fixed for future scalability.
  - **Good (4-6 points)**: The code is clean, well-structured, and modular. It follows basic coding best practices but could be optimised further for better performance and readability.
    - **4.1-4.5 points**: The code follows basic best practices and is well-structured and modular, but there is room for performance optimisation.
    - **4.6-5.0 points**: The code is functional and modular, but some areas could be better optimised for scalability and efficiency.

- - - **5.1-5.5 points**: The code is clean and well-organised with clear modularity, following most coding best practices, though there may be minor inefficiencies.
      - **5.6-6.0 points**: The code is well-structured, with good modularity and follows basic best practices. Some minor improvements could be made to optimise performance or readability.
  - **Very Good (6-8 points)**: The code is highly modular, efficient, and follows best practices. It is clean and easy to read, ensuring long-term maintainability.
    - **6.1-6.5 points**: The code is highly modular and follows best practices for efficiency. The structure is clear, but there may be minor performance optimisations needed.
    - **6.6-7.0 points**: The code is well-optimised, modular, and clean. There are a few inefficiencies, and the solution is ready for future improvements with minimal changes.
    - **7.1-7.5 points**: The code is highly modular, clean, and well-optimised. The structure follows best practices and is easy to maintain, though minor optimisation tweaks could enhance performance.
    - **7.6-8.0 points**: The code is efficient and modular, following best practices. It is well-organised, maintainable, and optimised for future scaling with minor refinements possible.
  - **Excellent (8-10 points)**: The code is elegant, highly optimised, and extremely modular. It follows advanced best practices for performance, scalability, and maintainability, ensuring future extensibility and easy updates.
    - **8.1-8.5 points**: The code is highly elegant and optimised. It follows advanced best practices for performance and scalability, but there may be small areas for improvement.
    - **8.6-9.0 points**: The code is exceptionally modular and optimised for performance, adhering to advanced best practices for both scalability and maintainability.
    - **9.1-9.5 points**: The code demonstrates industry-level performance and is extremely modular. It is perfectly optimised for scalability and future growth.
    - **9.6-10.0 points**: The code is exceptional in its elegance and optimisation. It follows cutting-edge practices for performance, scalability, and maintainability, ensuring future extensibility and easy updates.

- **Documentation and Explanation of Code**

- ○ **Poor (0-2 points)**: The code is poorly documented, with minimal or no comments. There is no clear explanation of the logic or the design choices.
    - ■ **0.0-0.5 points**: The code has no documentation or comments. There is no explanation of the logic or design, making the code difficult to understand or maintain.
    - ■ **0.6-1.0 points**: The code includes minimal documentation, with little comments or no explanations. It is unclear how the code works or why certain design choices were made.
    - ■ **1.1-1.5 points**: The code is poorly documented, with insufficient comments. Key sections of the code are unclear, making it difficult for others to follow or maintain the code.
    - ■ **1.6-2.0 points**: Some basic comments exist, but they are not detailed or specific enough to explain the logic or design decisions, making the code hard to follow.
- ○ **Fair (2-4 points)**: Some documentation is provided, but it is incomplete or lacks sufficient detail. Key sections of the code are not commented on or explained.
    - ■ **2.1-2.5 points**: The code contains basic comments but lacks sufficient detail or explanation for key sections. There is a lack of clarity in the logic and design decisions.
    - ■ **2.6-3.0 points**: Some parts of the code are commented, but important sections, such as some core functions and code logic, are missing explanations, making the code harder to follow.
    - ■ **3.1-3.5 points**: The documentation includes some explanations, but critical sections of the code are not commented on or are only briefly described.
    - ■ **3.6-4.0 points**: Some documentation is provided, but it is incomplete or lacks depth. Many important sections are either unexplained or under-explained.
- ○ **Good (4-6 points)**: The code is adequately documented with clear comments and explanations. The repository includes basic documentation for setup and use.
    - ■ **4.1-4.5 points**: The code is well-commented in most places, providing explanations for the key sections, though some areas could benefit from more detail or clarification.
    - ■ **4.6-5.0 points**: The code is adequately documented, with clear comments explaining the logic of most sections. The README file includes basic instructions for setup and usage but may not be comprehensive.
    - ■ **5.1-5.5 points**: The code contains clear and detailed comments on key sections. The README includes basic setup and usage instructions, though further explanation of more advanced features might be helpful.

- **5.6-6.0 points**: The code is well-documented, with clear comments for most functions and logic. The README provides sufficient setup and usage instructions but could benefit from further examples or explanations of complex features.
- **Very Good (6-8 points)**: The code is well-documented, with clear comments explaining key sections of the code. The README file is comprehensive and provides full instructions for installation, usage, and further development.
  - **6.1-6.5 points**: The code is well-documented, with detailed comments explaining key sections and the overall structure. The README includes comprehensive installation and usage instructions but could benefit from more detailed explanations or examples.
  - **6.6-7.0 points**: The documentation is extensive, with clear explanations for the logic behind key sections. The README is detailed, providing instructions for installation, usage, and some troubleshooting.
  - **7.1-7.5 points**: The code is well-commented, with clear explanations for all functions and major design decisions. The README is comprehensive, detailing installation, usage, and further development.
  - **7.6-8.0 points**: The code is thoroughly documented with clear explanations for every important section. The README is detailed, covering all necessary instructions for setup, usage, and contributing.
- **Excellent (8-10 points)**: The code is extensively documented with thorough, clear comments explaining every function, variable, and design decision. The README is exemplary, including detailed instructions, examples, and troubleshooting tips.
  - **8.1-8.5 points**: The code is extensively documented, with clear, detailed comments explaining every function and design decision. The README includes detailed instructions for setup, usage, and contributing, with real-world examples.
  - **8.6-9.0 points**: The code is exceptionally documented with thorough explanations for every part of the code. The README is exemplary, providing clear examples, setup instructions, and troubleshooting tips.
  - **9.1-9.5 points**: The documentation is exceptional, with clear and thorough comments for every function, variable, and design choice. The README is comprehensive, with examples and troubleshooting instructions.
  - **9.6-10.0 points**: The code is perfectly documented with comprehensive, clear comments for all aspects of the code. The

README is industry-level, with complete instructions, advanced examples, and troubleshooting guides.

4. Presentation and Pitching (20%)

- **Clarity and Structure**
  - **Poor (0-2 points)**: The presentation lacks clarity, is disorganised, and fails to communicate key points. It is hard to follow and misses critical information.
    - **0.0-0.5 points**: The presentation is disorganised and unclear, making it difficult to understand the solution. Key points are missing or not communicated effectively.
    - **0.6-1.0 points**: The presentation communicates some ideas but is disjointed and lacks clear structure. The message is unclear, and several key points are missing.
    - **1.1-1.5 points**: The presentation has a basic structure but lacks coherence or flow. The key ideas are communicated, but the overall narrative is weak, and some important information is missing.
    - **1.6-2.0 points**: The presentation is hard to follow, with poor transitions between points. Critical information is missing or poorly explained.
  - **Fair (2-4 points)**: The presentation communicates basic ideas but is somewhat disorganised, with areas lacking detail or clarity.
    - **2.1-2.5 points**: The presentation is somewhat structured, but some sections are unclear or lack detail. The key points are communicated, but the overall structure could be improved.
    - **2.6-3.0 points**: The presentation communicates basic ideas, but some sections could benefit from better organisation or more detailed explanations. It's mostly clear, but the flow of ideas is not smooth.
    - **3.1-3.5 points**: The presentation is generally clear, but the organisation could be improved. Some key points need further refinement to ensure they're effectively communicated.
    - **3.6-4.0 points**: The presentation conveys the main ideas clearly, but some sections may lack detail or could be better organised for a more seamless narrative.
  - **Good (4-6 points)**: The presentation is clear, well-structured, and communicates the main points effectively. Some areas could be more concise or refined.
    - **4.1-4.5 points**: The presentation is organised, clear, and communicates the main points well. However, some sections could be more concise or refined to improve clarity.
    - **4.6-5.0 points**: The presentation is well-structured, clearly communicating the problem, solution, and impact. Some minor improvements could make it more concise and impactful.

- - **5.1-5.5 points**: The presentation is clear and concise, effectively communicating the main ideas. The overall structure is strong, but some areas could be further refined for clarity or engagement.
    - **5.6-6.0 points**: The presentation is well-organised, effectively communicating the core message with clarity. It may benefit from minor tweaks to enhance engagement or refine the flow of information.
  - **Very Good (6-8 points)**: The presentation is engaging and well-structured, effectively communicating the problem, solution, and impact.
    - **6.1-6.5 points**: The presentation is engaging and well-structured, with clear communication of the problem, solution, and impact. It flows logically but could be enhanced with more compelling transitions.
    - **6.6-7.0 points**: The presentation is very clear, well-structured, and engaging. The narrative is compelling and well-paced, though minor adjustments could make it more impactful.
    - **7.1-7.5 points**: The presentation is clear, well-structured, and engaging. It clearly communicates the problem, solution, and impact in a compelling way.
    - **7.6-8.0 points**: The presentation is dynamic and engaging, with a strong narrative that effectively communicates the problem, solution, and impact. It's well-structured and persuasive.
  - **Excellent (8-10 points)**: The presentation is compelling and dynamic and seamlessly communicates the story. It is easy to follow, and the solution is presented in a persuasive, impactful way.
    - **8.1-8.5 points**: The presentation is compelling and dynamic, and clearly communicates the story. It's easy to follow and effectively persuades the audience of the value of the solution.
    - **8.6-9.0 points**: The presentation is highly dynamic, seamlessly communicating the solution with impact. The problem and solution are conveyed persuasively, and the narrative is smooth and compelling.
    - **9.1-9.5 points**: The presentation is exceptional—engaging, dynamic, and clear. It persuasively communicates the story, making the solution compelling and easy to understand.
    - **9.6-10.0 points**: The presentation is flawless, compelling, and dynamic, with a clear and impactful narrative that engages the audience and effectively communicates the solution and its value.

- **Visual Appeal and Engagement**

- ○ **Poor (0-2 points)**: The slides are cluttered, poorly designed, and fail to engage the audience.
  - ■ **0.0-0.5 points**: The slides are disorganised, cluttered, and difficult to read. They fail to engage the audience or support the narrative.
  - ■ **0.6-1.0 points**: The slides are poorly designed, with overcrowded content or inconsistent visuals, making them difficult to follow and lacking engagement.
  - ■ **1.1-1.5 points**: The slides are functional, but the design is unappealing and does not support the presentation well. They do not engage or captivate the audience.
  - ■ **1.6-2.0 points**: The slides are basic and functional but lack a polished design. They may fail to support the message or engage the audience effectively.
- ○ **Fair (2-4 points)**: The slides are functional but not particularly engaging. They could benefit from a cleaner design and more engaging visuals.
  - ■ **2.1-2.5 points**: The slides are functional but lack creativity. They are difficult to read or poorly formatted, making it hard for the audience to follow.
  - ■ **2.6-3.0 points**: The slides are clean, but the design is unappealing or lacks visual engagement. They support the narrative, but they fail to engage the audience fully.
  - ■ **3.1-3.5 points**: The slides are functional, with basic design elements that support the presentation. However, they could be more visually appealing to improve audience engagement.
  - ■ **3.6-4.0 points**: The slides are adequate, but the design is basic or lacks impact. They could benefit from simplification or more engaging visuals.
- ○ **Good (4-6 points)**: The slides are well-designed and visually appealing, enhancing the presentation. They engage the audience and support the narrative.
  - ■ **4.1-4.5 points**: The slides are visually appealing, with a clean design that supports the presentation. Some slides could benefit from more engaging visuals to enhance the narrative.
  - ■ **4.6-5.0 points**: The slides are well-designed, clean, and visually engaging, but they could be further optimised to strengthen the narrative or increase impact.
  - ■ **5.1-5.5 points**: The slides are well-designed and engaging, with a professional layout that complements the narrative, though some slides could be simplified or refined for clarity.
  - ■ **5.6-6.0 points**: The slides are polished and visually striking, effectively supporting the message and engaging the audience throughout the presentation.

- **Very Good (6-8 points)**: The slides are polished, professional, and visually striking, supporting the pitch with clarity and engagement.
    - **6.1-6.5 points**: The slides are well-designed, polished, and visually striking. They effectively engage the audience and support the narrative with clarity and impact.
    - **6.6-7.0 points**: The slides are professional and dynamic, with visuals that enhance the storytelling. They are visually striking and complement the message very well.
    - **7.1-7.5 points**: The slides are visually stunning and well-integrated into the presentation. They enhance the story and are highly engaging.
    - **7.6-8.0 points**: The slides are exceptionally polished and visually captivating, supporting the narrative with clarity and providing an impactful experience for the audience.
- **Excellent (8-10 points)**: The slides are exceptional, visually stunning, and creatively designed. They captivate the audience and reinforce the message in a compelling way.
    - **8.1-8.5 points**: The slides are exceptional, visually stunning, creative, and perfectly aligned with the message. They captivate the audience and reinforce the solution in a compelling way.
    - **8.6-9.0 points**: The slides are flawlessly designed, highly engaging, and creative. They are exceptionally impactful, complementing and reinforcing the pitch's main points.
    - **9.1-9.5 points**: The slides are industry-level, exceptionally creative, and perfectly aligned with the presentation. They are highly engaging and support the narrative with maximum impact.
    - **9.6-10.0 points**: The slides are exceptionally creative, flawlessly designed, and impactful, captivating the audience and reinforcing the message in a dynamic and compelling way.