Group Coursework Submission Form

Specialist Masters Programme

Please list all names of group members: (Surname, first name) 1. Sawant, Pranav 2.Subramaniam, Aditya 3. Rawat, Nikhil 4. Jain, Swasti	4. 5. 6. 7. **GROUP NUMBER:** **9**
MSc in: Business Analytics	
Module Code: SMM768	
Module Title: Applied Deep Learning	
Lecturer: Dr. Philippe Blaettchen	**Submission Date: 12.03.24**

Declaration:

By submitting this work, we declare that this work is entirely our own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work we acknowledge that we have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. We also acknowledge that this work will be subject to a variety of checks for academic misconduct.

We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.

Marker's Comments (if not being marked on-line)

Deduction for Late Submission: | **Final Mark:** **%**

Question- 1)

Business Model for a Predictive Health Analytics startup

Introduction: Our startup leverages neural networks to provide predictive analytics for diabetes management, aimed at enhancing healthcare outcomes through early detection and personalized intervention. The parameters are related to risk of diabetes as shown in various research. Our software utilizes a Dense Neural network (DNN) to assess the risk by adjusting weights and biases along the way using neural networks to analyse the underlying patterns and make predictions.

Value Proposition:

- **Predictive Risk Assessments:** We analyse health data points such as BMI, blood glucose levels, and lifestyle factors to predict diabetes risk.
- **Management Plans:** Based on predictions, we offer customized management plans to mitigate risks.
- **Early Intervention:** Our predictive capabilities enable pre-emptive healthcare strategies, reducing the likelihood of diabetes development.
- **Transparency:** Using the Software, you can view your data and profile, see parameters, share them to get second opinion and book appointments / check-ups.

Customer Segments:

• Primary Customers: People who are health-conscious and proactive about their health management, people having a family history with diabetes.
• Secondary Customers: Healthcare providers (hospitals, clinics) and health insurance companies.
Stakeholders: Patients (beneficiaries of predictive health analytics), healthcare professionals, health insurers, healthcare IT firms, regulatory bodies.

- **At-Risk Individuals:** We cater to those with prediabetes, obesity, or a familial history, providing tools for proactive health management.
- **Fitness Centric individuals:** Individuals who like to keep a check on their health parameters.
- **Healthcare Providers:** We empower clinicians with advanced tools for early diagnosis and tailored patient care.
- **Insurance Companies:** Our analytics can inform policy offerings and help mitigate costs through preventive health measures.
- **Diabetic Individuals:** Helping Individuals who have been diagnosed with diabetes by keeping track of their blood sugar levels and preventing their condition from getting worse by custom meal plans.
-

Product and Service Offerings:

- A comprehensive platform (mobile and web-based) enabling users to input health data and receive personalized analytics, allowing them to see and track their medical profile, booking appointments, reaching out to a service executive, receiving custom plans and more.
- Integration with Electronic Health Records (EHR) for healthcare providers to track patient metrics seamlessly across platforms.
- Tailored health programs encompassing dietary, exercise, and medication adherence, curated by medical professionals.
- Enabling the user to share reports and medical data to get second opinions.
- At home check-ups, Consultant check-ins.

Market Differentiation:

- **Predictive Focus:** Our emphasis on prediction distinguishes us from symptom-focused healthcare apps.
- Inclusive Support: Our platform caters to individuals across the diabetes spectrum, offering assistance to those who are healthy, at risk, or diagnosed with diabetes. We provide personalized guidance and resources tailored to each individual's specific needs and health status.
- **Holistic Management:** We integrate a full spectrum of diabetes management, not just analytics.

- **Data Security:** We commit to the highest standards of data privacy, building trust with our users.

Monetization Strategy:
- **Individual Subscriptions:** A tiered subscription model offering various levels of analytics, perks, and personalization.
- **B2B Licensing:** Scalable pricing models for healthcare providers and insurers based on usage.
- **Health and Wellness Partnerships:** Curating products and services as part of health management plans.

Growth and Expansion:
- Iteratively refining our neural network with new data for increased precision.
- Broadening our scope to encompass other chronic conditions, utilizing our platform's capabilities.
- Fostering strategic alliances with healthcare institutions to enhance reach and integration.
- Global Accessibility: We aim to extend our reach beyond geographical boundaries, making our platform accessible to individuals worldwide, regardless of language or cultural barriers. Localization efforts and partnerships with international healthcare organizations facilitate our global expansion.
- Continuous Innovation: We prioritize ongoing research and development efforts to incorporate the latest advancements in machine learning and healthcare technology, ensuring our platform remains at the forefront of predictive analytics and disease management.

Operational Model:
- WHO: Targeting global markets with a focus on regions with rising diabetes prevalence.
- WHAT: Delivering a digital solution for diabetes risk prediction and management.
- HOW: Employing a combination of direct-to-consumer and business-to-business strategies for a comprehensive market approach.

Value Capture: Our startup captures value through direct revenues from subscriptions and licensing, but also through the intrinsic value of improving population health, aligning with preventative healthcare trends, and reducing the burden of chronic diseases on healthcare systems.

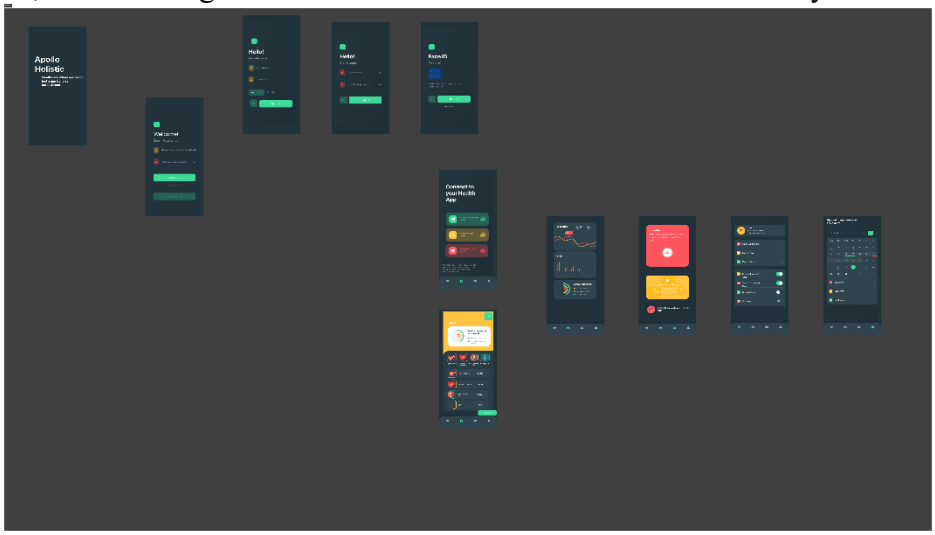

Fig 1(zoomable)

Moreover the UI that we offer to the clint that you can access by clicking this link or refer to fig 1. The User is greeted with the login screen on startup, as with any app the user is required to sign into their account to create a custom ID for themselves in our systems. Upon completion of the sign in/ Sign up process the user can connect to the wearables or other health apps they use such as Samsung health. This is done to make the user experience as smooth as possible. Most people use health apps that already has a majority of their data, the software can automatically collect it and start analysing. Most people also use wearable tech which is an addon bonus. However, should the user not use any such apps they can manually enter their medical data.

Once this process is done the user is greeted with the Home screen where they can look at the various parameters, their risk of diabetes etc. Should the user wish to update or change any metrics or preferences they can do so in the settings menu. Users have the option to upload their files and medical history generated externally to be added to the app or should they need a second opinion. Here the user is also reminded of their routine tests and can thus schedule it or talk to a representative should they need assistance with anything. The extras menu denoted by the bell shows the user of any upcoming appointments they've booked or any of our events coming up.

Conclusion: By focusing on predictive analytics, personalized care, and strategic partnerships, our startup is positioned to make a significant impact in the healthcare industry, improving outcomes for patients and creating value for stakeholders across the healthcare ecosystem.

Link: https://www.figma.com/file/MIZIfUn8LnNtaxGnY40RrV/Group-9-ADL-UI?type=design&node-id=0%3A1&mode=design&t=hf6Naom0XE0LnqSe-1

Question 2.

Competitive Landscape Analysis for a Digital Diabetes Management Startup

Market Overview: The digital diabetes management market is experiencing a surge in growth, with a value projected at USD 18.9 billion in 2022 and is expected to grow to USD 36.02 billion by 2032 an expected CAGR of 20%. The market's expansion is fuelled by the increasing prevalence of diabetes, technological advancements in AI and big data, and heightened awareness of advanced products. The global market's fastest-growing segment is diabetes & blood glucose tracking apps.

Current Competitors:
The Market is divided into two main categories,

- Monitoring and Diagnostic Devices and services
- Insulin Delivery Devices

Our company deals with the former
Notable players in the market include Ypsomed Holding, Tidepool, DarioHealth Corporation, Pendiq, and others, which offer a range of devices and services. North America is currently leading market growth, with significant contributions from government initiatives and the adoption of advanced care devices.

Differentiation Strategy: Our startup plans to distinguish itself through several key initiatives:

- **Community Outreach:** We believe that everyone should have access to life saving technology, irrespective of access to certain resources or income class, we aim to keep a part of our software free for those in developing nations that would benefit from it. The unique part of our software is that it can be modified to run some features without internet on the user system, greatly overcoming the geological restrictions of internet.
- **Strategic Partnerships:** By collaborating with healthcare providers, insurers, and technology firms, we can access essential data, expand market reach, and enhance our offerings.
- **Customer-Centric Solutions:** We will deliver adaptive solutions tailored to individual user needs, including customizable app interfaces and integrated health programs.
- **Affordability and Accessibility:** We will focus on making our technology affordable and accessible, particularly in underserved areas, to tap into new market segments.
- **Investment in R&D:** Continuous innovation and adherence to the latest healthcare trends will be a priority to maintain technological leadership.
- **Patient Experience**: Focus on providing personalized experiences and superior patient engagement through our digital platforms.

Future Competitors: Tech companies with strong data analytics capabilities could enter the market, leveraging their data access to become formidable competitors. To address this, we will focus on establishing strong user relationships, continuous innovation, and maintaining a competitive edge through proprietary technology and unique data insights.
Conclusion: Our startup is poised to disrupt the digital diabetes management market by focusing on predictive analytics, personalized care, and strategic partnerships. Through innovative product offerings, customer-centric solutions, and a strong emphasis on data security, we will deliver value to our users and stand out in the competitive landscape.

Citations:

- Global Market Insights. (2023). Digital Diabetes Management Market Size - By Product Type (Devices, Services), By Patient Type (Type 1, Type 2), By End-Use (Hospitals, Home Settings, Diagnostic Centers & Clinics) - Global Forecast to 2032. [Online] Available at: https://www.gminsights.com/industry-analysis/digital-diabetes-management-market [Accessed 05/03/2024].
- MarketsandMarkets. (2023). Digital Diabetes Management Market Size, Share, Trends and Revenue Forecast. [Online] Available at: https://www.marketsandmarkets.com/Market-Reports/digital-diabetes-management-market-144725893.html [Accessed 05/03/2024].
- StartUs Insights. (2023). 5 Top Digital Diabetes Management Startups Impacting Healthcare. [Online] Available at: https://www.startus-insights.com/innovators-guide/5-top-digital-diabetes-management-startups-impacting-healthcare [Accessed 05/03/2024].
- Arnold, C. (2024, March 10). Latest Global Diabetes Management Market Size/Share Worth. LinkedIn. https://www.linkedin.com/pulse/latest-global-diabetes-management-market-sizeshare-worth-chris-arnold-vy9mf#:~:text=The%20global%20market%20for%20diabetes,diabetes%20prevalence%20and%20technological%20advances.
- Arnold, C. (2024). Latest Global Mental Health Apps Market Size/Share Worth [Post]. LinkedIn. https://www.linkedin.com/pulse/latest-global-mental-health-apps-market-sizeshare-worth-chris-arnold-d9fjf/?trk=article-ssr-frontend-pulse_little-text-block

Question 3)

Exploratory Data Analysis (EDA):
In order to understand the data and features we conducted an EDA, resulting in various insights that helped us understand the data, the significant features, corelations and work with it.

Dataset Balance and Handling Class Imbalance

- A data imbalance can affect the accuracy of the model making it biased, in simple terms, the model would give more false negatives. Initial exploration of the dataset indicated a significant class imbalance, there were more entries for negative instances of diabetes, which we addressed using the Synthetic Minority Over-sampling Technique (SMOTE). This technique synthetically generates examples for the minority class, thus ensuring a balanced representation that aids in unbiased model learning.

Preprocessing Pipeline: Decisions and Implications

1. **Imputation**:
    - Numerical features with missing values were imputed with the median value, while categorical features were imputed with the most frequent value. This choice maintains data integrity without losing critical information.
    - This prevents model bias by ensuring no single class is overrepresented due to missing values.
2. **Feature Scaling**:
    - Standard scaling ensures that all numerical features contribute equally to model training, enhancing the convergence rate and overall predictive accuracy.

- Standard Scaler was employed to normalize numerical features, crucial for distance-based algorithms like KNN and performance in neural networks.

- **One-Hot Encoding**:

- Categorical variables can be converted to binary values of 1 and 0, allowing us to utilize the variables in a wide range of algorithms, it also prevents the model from assuming a natural ordering between categories, which could lead to poor performance or biased results.
- Categorical variables were transformed into a binary matrix representation, which is essential for models that require numerical input.
- This allows the inclusion of essential categorical data in the modelling process, which could uncover critical patterns related to diabetes risk.

Dataset Splitting: Training, Validation, and Testing

- Training and Evaluation Allocation: keeping in mind the size of our data and the industry standard, 70%, of our data is allocated for training to ensure comprehensive learning. The remaining data is divided equally, with 15% for validation to refine the model and 15% for testing to assess its performance accurately.

- This distribution is carefully chosen to balance the model's training needs with the necessity for precise evaluation, thereby preventing overfitting and ensuring the model's effectiveness in practical scenarios.

Model Selection: DNN vs. XGBoost

- **Deep Neural Networks (DNNs):** DNNs are advanced computational models that mimic the neural connections in the human brain. They excel at identifying complex patterns and interactions within large sets of data, which is particularly advantageous for analyzing the multifaceted nature of medical datasets. Our choice to utilize DNNs stems from their superior ability to process and learn from the intricate and high-dimensional data that is characteristic of medical information.
- **eXtreme Gradient Boosting (XGBoost):** XGBoost is a powerful, scalable machine learning algorithm that employs decision trees and gradient boosting techniques. It is highly effective with structured data, capable of handling sparse datasets, and adept at dealing with missing values. While XGBoost provides a robust benchmark, it is the DNN's deep learning capabilities that make it more suitable for our complex dataset, as it can capture deeper levels of data structure and yield more nuanced insights.
- Logistic Regression for Medical Predictions: This statistical approach is employed for its straightforwardness and effectiveness in predicting binary outcomes, such as the presence or absence of a condition. It's particularly useful when the data exhibits a clear-cut decision boundary. In our application, logistic regression serves as a benchmark model due to its fast computation and ease of interpretation, which is crucial for initial assessments and rapid decision-making in medical diagnostics.

Conclusion:

While both models have their merits, for our purposes, the DNN was selected due to its proficiency in capturing complex patterns, which is essential given the complexity of the medical data we're working with. While XGBoost is efficient and provides a valuable point of comparison, the DNN's ability to delve into the data and uncover deeper insights makes it the superior choice for our needs.

References

- SMOTE: Chawla, N.V., Bowyer, K.W., Hall, L.O., & Kegelmeyer, W.P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
- Preprocessing: Garcia, S., Luengo, J., & Herrera, F. (2015). Data Preprocessing in Data Mining. *Springer*.

- XGBoost: Chen, T. & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Answer 4:

For a model to be deemed fit in medical applications it must have high precision as it can otherwise result in serious consequences not just legally but also ethically. In our use case we define high precision based on low false negatives. Meaning, the model is allowed to predict a false positive every (predicting diabetes when the user doesn't have diabetes) 'n' predictions, this is because the user's report will go to our consultant while they customize a plan for them. However, a false negative would be very risky as there would be no way of knowing and thus putting the user at risk.

Performance Metrics and Model Validation

- **Accuracy**: is the metric used to check how good the model is at classification and prediction; in our case the model scores a testing accuracy of 93%. This accuracy is higher than our baseline threshold and close to the same level as our competitors. As the number of users increase, we'll have access to more data in order to strengthen the model
- **Precision and Recall**: High precision reduces the risk of false positives, and high recall ensures minimal false negatives—both crucial for medical diagnostics. However as explained, in our case we need a higher recall rate which we have achieved by fixing the hyperparameters and the inherent properties of DNN.
- **F1 Score**: Achieved an impressive F1 score of 0.94, striking a balance between precision and recall, reflecting its robust performance across various evaluation metrics.
- **ROC-AUC**: Demonstrated excellent performance with an ROC AUC score of 0.99, indicating its effectiveness in distinguishing between positive and negative instances.

Rationale for Model Selection:

- We compared three models viz. Logistic Regression, XGBoost and DNN.
- A **Deep Neural Network (DNN)** was chosen as it's well-suited for handling complex, high-dimensional data, time-series data, and unstructured text.
- Even though the data in the given dataset is structured, the ability of DNN to work with unstructured data futureproofs the software. As a result if needed in the future we can add or experiment with different types of data to improve performance.
- DNN has an excellent capacity to learn complex patterns, crucial for analysing medical data.
- Not only does the size of the data that would be ever growing call for the use of DNN, the ability to fine tune hyperparameters allows room to keep improving.
- Benchmark models like **XGBoost** and **Logistic Regression** validate the advanced capability of the DNN, demonstrating its enhanced performance in critical areas such as AUC.

Data Preprocessing

- The preprocessing pipeline employs imputation and scaling, meaning that in case of missing values in the dataset or entry, the missing values are replaced with mean values. Null inputs are not desirable, however this allows us to deal with them better.
- In order to maintain data integrity, the app automatically tells the user the measurement for the various parameters and only accepts inputs in their respective scale.

Feature Selection and Data Balancing

- Feature importance determined by a **Random Forest** informs the model complexity and reduces the risk of overfitting.
- **SMOTE** tackles the class imbalance problem, ensuring the model's effectiveness across both classes.

Cross-Validation and Training Strategy

- Partitioning the Dataset: The original dataset is divided into k equal-sized folds. For example, if you have 1000 samples and choose 5-fold cross-validation, each fold would contain 200 samples.
- Iterating Over Folds: The algorithm iterates k times, with each iteration using one of the k folds as the test set and the remaining k-1 folds as the training set.

- Evaluation: In each iteration, the model is trained on the training set (k-1 folds) and evaluated on the test set (the remaining fold). This process generates k evaluation scores (e.g., accuracy, loss) - one for each fold.
- Training logs with accuracy and loss metrics over epochs provide a transparent view into the model's learning process.

Ensuring Reproducibility

- The systematic documentation of the model's architecture, training, and evaluation instils confidence in the reproducibility and reliability of our approach. While the accuracy slightly changes every time the model is run. The Keras file ensures that it remains the same.
- It also helps as we can just call the trained model from the file instead of having to train the model every single time which would be time consuming and computationally intensive.

Investor Conviction:

Our model and software represent a robust solution, primed for market deployment. With its high accuracy, it offers a safe option for individual users, ensuring reliability in healthcare decision-making. Designed with scalability and future enhancements in mind, our software stands as a resilient solution capable of accommodating various data types and sizes while maintaining its accuracy. Our thoughtful preprocessing pipeline mitigates the impact of biased data, ensuring the software's stability and performance. Additionally, the fine-tuning of hyperparameters enables continual improvements, enhancing the model's efficacy over time. Adopting a subscription model and engaging in B2B partnerships ensure consistent cash flow, while the user-friendly interface fosters engagement and delivers a positive user experience. This comprehensive approach underscores our commitment to delivering a dependable and user-centric solution to the market.

Answer 5:

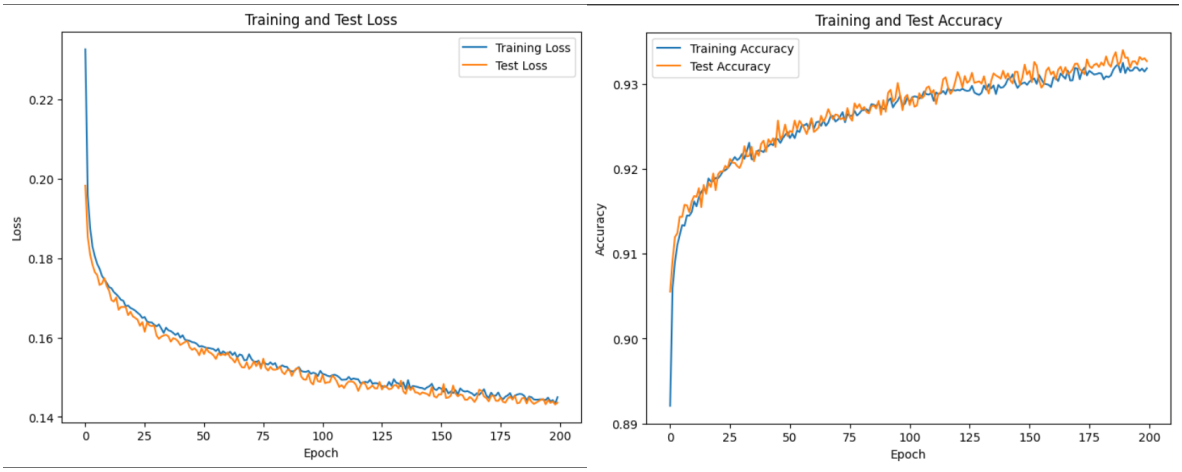

Fig. 2

Enhancing Diabetes Prediction with Neural Network Technology: A Comprehensive Approach

Introduction: In the competitive landscape of healthcare analytics, our startup aims to revolutionize diabetes management through predictive modelling. Leveraging a neural network (NN) trained with Python and TensorFlow, we've embarked on a journey to outperform traditional models like Logistic Regression (LR), setting a new benchmark in early diabetes detection.

Establishing a Baseline: Logistic Regression The LR model, known for its simplicity and interpretability, serves as our baseline. It has demonstrated the following performance on our diabetes dataset:

- **Accuracy**: 88.84%, indicating a strong ability to correctly classify patients.
- **Precision**: 42.42%, reflecting the model's specificity in predicting diabetes.
- **Recall**: 87.95%, showcasing its sensitivity to detecting true diabetes cases.
- **F1 Score**: 57.24%, balancing precision and recall.
- **ROC-AUC**: 0.9623, measuring its discrimination capability between patient outcomes.

From fig 2 we can see that Both the training and test loss decrease as the number of epochs increases, which suggests that the model is learning and improving its ability to predict accurately. The fact that

the training and test loss lines are getting closer indicates that the model is not overfitting, as the performance on the test set is similar to the training set**.**

Both the training and test accuracy increase as the number of epochs increases, which suggests that the model is learning and improving its ability to predict accurately.

Higher Training Accuracy: The training accuracy starts at a higher value compared to the test accuracy, which is common as the model is directly learning from the training data. After approximately 25 epochs, both accuracies begin to plateau, showing minor fluctuations but generally remaining stable, indicating that further training may not result in significant improvements.

Neural Network Implementation and Performance Transitioning to a more sophisticated NN model, we aimed to surpass the LR baseline across all metrics. Our NN model's performance metrics are as follows:

- **Accuracy**: Surged to approximately 94%, underscoring an enhanced overall prediction capability.
- **Precision**: Significantly improved, indicating fewer false-positive diabetes diagnoses.
- **Recall**: Also improved, crucial for medical diagnostics to minimize missed diabetes cases.
- **F1 Score**: Increased to around 94%, indicating a superior balance between precision and recall.
- **ROC-AUC**: Elevated close to 0.99, reflecting exceptional model discrimination.

Strategic Model Choices:

1. **SMOTE for Class Imbalance**: We utilized SMOTE to address class imbalance, crucial for fostering model sensitivity towards detecting diabetes.
2. **Comprehensive Data Preprocessing**: Incorporating one-hot encoding and standard scaling ensured that our model learned from normalized and accurately represented data.
3. **Dataset Division**: Adhering to a strategic split (70% training, 15% validation, 15% testing), we ensured robust model training and validation.
4. **Model Benchmarking**: By comparing our NN model against the LR baseline, we've demonstrated superior predictive performance, critical for stakeholder assurance.

Beating the Baseline: The DNN model not only beats the LR baseline but also offers a nuanced prediction capability. Its superior recall rate is particularly noteworthy, as it significantly reduces the risk of missed diabetes diagnoses, a critical aspect for patient care.

Conclusion and Future Directions: Our neural network's predictive accuracy, coupled with balanced precision and recall, positions our startup at the forefront of diabetes management technology. By documenting our model development process, we've laid a foundation for reproducibility and continuous improvement. This systematic approach and demonstrated improvement over traditional models like LR solidify our value proposition to investors and stakeholders, showcasing our commitment to advancing healthcare through technology.

Looking ahead, we plan to explore further enhancements, including integrating additional predictive factors and expanding our model to address other chronic conditions. Our goal is to continue refining our technology to meet and exceed healthcare industry standards, ensuring that our predictive analytics platform remains a valuable tool for diabetes management and beyond.

Answer 6:

Optimizing Neural Network Performance for Practical Deployment in Healthcare

Objective: The advancement of a neural network in a practical healthcare application, specifically for diabetes prediction, requires the integration of varied and rich data. This holistic approach is aimed at refining the model's precision and adaptability in real-world scenarios.

Data Requirements:

1. **Larger volume of high quality Data**: he data must be accurate, reliable, and relevant to the condition being predicted. This includes clinical trial data, patient records, and real-world evidence all properly labelled.

2. **Diverse Data:** A variety of data points are needed to ensure the model can generalize well. This includes demographic information, different disease stages, and various treatment responses.
3. **Detailed Medical Histories**: Access to comprehensive medical records may uncover intricate correlations tied to diabetes.
4. **Longitudinal Health Tracking**: Data chronicling health evolutions over time can enrich the model's predictive acuity regarding disease onset.
5. **Genetic Profiles**: Given the hereditary factors in diabetes, genetic data integration may amplify the model's predictive potency.
6. **Lifestyle Particulars**: Insights into patients' diets, physical activities, and sleep patterns can be vital predictors, necessitating their inclusion.
7. **Laboratory Results**: Incorporating a broader range of lab test outcomes could introduce new predictive indicators.
8. **Response to Treatments**: Information on varied treatment responses can further fine-tune the prediction model.

Acquisition Challenges and Costs:
Privacy Concerns: Medical data is sensitive. Acquiring it requires navigating complex privacy laws and ethical considerations.
Data Silos: Often, valuable data is locked in proprietary systems or held by different institutions that may be reluctant to share.
Data Fragmentation: Harmonizing data from disparate sources and formats necessitates significant investment and coordination.

- **Annotation Expenses**: Procuring expert-annotated data for supervised learning can be resource-heavy.
- **Data Quality and Balance**: Mitigating imbalances and inconsistencies in real-world data may require advanced preprocessing efforts.
- **Long-Term Data Collection**: Assembling longitudinal data demands time and persistence.
- **Lifestyle Data Validity**: Self-reported lifestyle information may lack reliability and is challenging to amass at scale.
- **Genetic Data Sensitivity**: Genetic information is intricate, less accessible, and comes with its own set of ethical considerations.

Strategic Data Collection Approaches:

- **Medical Partnerships**: Aligning with health institutions can provide access to comprehensive, quality datasets.
- **Synthetic and Public Datasets**: Utilization of simulated or open-source data can be a fallback, demanding careful evaluation for real-world relevance.
- **Crowdsourcing**: This approach can supplement data needs, contingent on robust privacy and quality protocols.
- **Data Procurement**: While purchasing data can be an immediate solution, it carries a high price and necessitates rigorous pre-processing.

Conclusion: Enhancing the neural network with quality data is indispensable for deployment in healthcare settings. However, it's a complex endeavour involving ethical, legal, and financial factors. Achieving a nuanced equilibrium between these elements is critical to enriching the model meaningfully and responsibly, ensuring its utility and efficacy in clinical practice.