RATIONALE:

Neonatal care is frequently neglected as hospitals struggle to provide proper assistance during this critical period for the child. Oftentimes, focus is directed toward immediate post-birth care, leaving new mothers with minimal support and a lack of medical guidance in the aftermath of childbirth. In order to help mothers smoothly transition from neonatal to postpartum care, we must create an AI model that identifies gaps in growth and potential genetic disorders.

ENGINEERING GOAL(S):

The objective of NeoCare is to develop a precise and dependable predictive model. This model will analyze key neonatal measurements and compare them against CDC growth charts to identify potential neonatal syndromes.

A significant focus will be placed on ensuring the model’s alignment with CDC standards for accuracy in syndrome identification. Additionally, the project should emphasize the importance of effective data handling and normalization. This should entail meticulously collecting and standardizing CDC data to ensure consistency and accuracy, a crucial step for reliable model functionality.

In the realm of algorithm development and refinement, the goal is to create robust algorithms within the model capable of accurately calculating the likelihood of various syndromes. We will employ visualization techniques to provide clear and understandable representations of the model’s predictions.

The project should also prioritize addressing and rectifying initial oversights, such as incorporating familial traits into head circumference analysis. This will enhance the model’s overall precision. We should also maintain thorough documentation to ensure the experiment’s replicability and transparency. The project will set out to establish a framework for the model’s validation and improvement. This includes updating it with a broader and more diverse dataset and aligning it with the latest CDC data and medical insights, which will solidify its effects on neonatal healthcare.

EXPECTED OUTCOMES:

The expected outcome is to create a predictive model to diagnose different genetic diseases in newborns to help new mothers understand the symptoms and treatment of their child’s condition. Using the data samples of newborn genetics, we will be able to test the accuracy of our model.

HYPOTHESIS:

If we create an AI model to individualize postpartum care, then the general health of the mother and their respective newborn will improve because by addressing and explaining the physical factors that affect the health of the baby and mother, the families will be able to understand the medical situation of the child properly; achieving this will relieve stress from the mother and lead to more effective and efficient care for the child.

RESEARCH QUESTION(GRQ):

How do mothers get specialized postpartum care for their child and themselves?

RISKS AND SAFETY:

One of the most prominent risks of our experiment is the possible misdiagnosis of a syndrome pertaining to a data set. If a misdiagnosis occurs, the condition of a young child may worsen. In some situations, permanent damage to an infant may occur if they are treated incorrectly. In order to mitigate a misdiagnosis, we intend on doing multiple trials in order to find errors or gaps in the predictive model.

PROCEDURES:

The materials required for this experiment include CDC Growth Charts and Data Tables for both male and female babies, a Python programming environment (or similar software capable of data analysis and visualization), and a sample dataset containing children's measurements such as Age, Length, Weight, Sex, and Head Circumference.

The first step we will need to take in order to begin is researching disorders associated with various growth abnormalities. This will be crucial in order to provide a pool of possible syndromes for us to use when creating our predictive model. We will research genetic disorders pertaining to the height, weight, and other important measurements in infancy.

After our research, we will narrow down the scope to the three (3) genetic disorders that we would like to focus on throughout the building of our model. These genetic disorders may be decided by factors such as the frequency in a population, the recognizable signs in children, and the risks of late diagnosis.

We will begin by collecting CDC growth data specific to prominent measurements of an infant and prepare a diverse dataset representing various children's measurements. This dataset will include crucial data points in infancy such as age, length, weight, sex, and head circumference.

We will then develop a Python script or utilize a similar data processing tool to input the child's measurements. The script should use logic to compare these measurements against CDC data. This should include major deviations in the trends of given graphs.

Upon developing the code, we will implement a scoring system to quantify the likelihood of each syndrome. For each syndrome, the individual scores must be calculated based on the deviation from CDC's normative data, with each score being capped between 0 and 100.

To test the model, we will use a prepared sample dataset to predict potential syndromes. We will record the flags, which should indicate potential syndromes, and calculate a score (using the scoring system) for each child in the dataset.

We will then create visual representations of the model's predictions using data visualization tools. This may include heatmaps, bar charts, histograms, and other methods of displaying data.

In order to understand the results, we will analyze the visualized data to interpret the predictions made by the model. We can compare the results against CDC's expected norms to evaluate the accuracy and reliability of the model's predictions.

It is imperative that we document every step in the procedure. This documentation will include details of the model development, testing, and analysis processes to ensure clarity and sufficiency for replication.

Finally, we will review the entire process to identify any potential improvements to be made. Based on this review, we will refine the model and the procedure (if necessary) to enhance its effectiveness and accuracy.

This structured procedure provides a comprehensive guideline for conducting the experiment, ensuring reproducibility in the future.

DATA ANALYSIS:

In our experiment to predict neonatal syndromes, we will analyze the data using a combination of statistical methods and data visualization techniques. The model will be designed to flag potential cases of the three (3) syndromes that we select. For each case, we will calculate a NeoScore – a numerical value representing the likelihood of each syndrome.

Once we have the scores and flags for each case, we will utilize data visualization tools to interpret these results effectively. We may create heatmaps and stacked bar charts to provide a clear view of the distribution and intensity of NeoScores across different syndromes.

The last step of our data analysis will likely involve a form of comparative review of the model’s predictions against the CDC's normative data. This will be key for evaluating the accuracy and reliability of our model. We will also carefully document and review each step of our analysis to ensure reproducibility and to identify any areas for potential improvement in future studies.

This comprehensive analysis approach should allow us not only to predict potential neonatal syndromes but also to quantify the likelihood of these conditions in a way that is both scientifically rigorous and visually interpretable.

Initially, our approach involved collecting and digitizing CDC data. A deviation occurred in the early phase of data collection, where the data normalization process was overlooked, leading to inconsistencies in the dataset. This issue was later identified and rectified by standardizing the data to a uniform format, ensuring accurate comparisons against the CDC benchmarks.

Additionally, the model development in Python encountered a minor setback due to an initial oversight in the algorithm that calculates the NeoScores. The calculation inaccurately handled edge cases where children's measurements were precisely on the percentile boundaries. This was later corrected by refining the logic to include boundary conditions, thereby improving the model's accuracy.

PROJECT SUMMARY:

During data analysis, the use of heatmaps and stacked bar charts effectively visualized the results. However, an initial mistake was made in the color coding of the heatmap, which led to some confusion in interpreting the likelihood of the syndrome. This was mitigated by revising the color scheme for more precise differentiation between the syndromes.

A notable deviation occurred in the interpretation phase. The initial analysis did not account for the possible impact of familial traits on the measurements, especially for head circumference in the context of microcephaly and macrocephaly. Recognizing this, we incorporated a review step to consider familial data, which provided a more nuanced understanding of the results.

Despite these deviations and initial mistakes, the project successfully achieved its objective. The final model demonstrated a robust ability to predict neonatal syndromes with a quantifiable degree of likelihood. Key improvements made during the project include data normalization, refinement of score calculations, enhanced visualization techniques, and the inclusion of familial trait considerations.

Future iterations of the project could benefit from integrating a more extensive and more diverse dataset, including genetic factors, to enhance the model's predictive capabilities further. Continuous validation against updated CDC data and clinical inputs will also be crucial for maintaining the model's relevance and accuracy.

BIBLIOGRAPHY:

Boston Children’s Hospital. (2023, February). *Growth problems*. Growth Problems | Boston Children’s Hospital. http://www.childrenshospital.org/conditions/growth-problems#:~:text=Growth%20problems%20may%20be%20a,growth%20hormone%20deficiency.

Wroblewski, M. E., Badik, C., & Bevington, J. (2015, September 1). *Head Growth*. Publications.aap.org.

https://publications.aap.org/pediatricsinreview/article-abstract/36/9/426/34910/Head-Growth?redirectedFrom=fulltext

Robert H. Shmerling, M. (2023, May 5). *How useful is the body mass index (BMI)?*. Harvard Health. https://www.health.harvard.edu/blog/how-useful-is-the-body-mass-index-bmi-201603309339

Safer-Healthier-People. (2000, May). Overview of the CDC growth charts. https://www.cdc.gov/nccdphp/dnpa/growthcharts/training/modules/module2/text/module2print.pdf