

# Modello Statistico per la Previsione del Peso Neonatale

## Contesto Aziendale

**Azienda:** Neonatal Health Solutions

**Obiettivo:** Creare un modello statistico in grado di prevedere con precisione il peso dei neonati alla nascita, basandosi su variabili cliniche raccolte da tre ospedali. Il progetto mira a migliorare la gestione delle gravidanze ad alto rischio, ottimizzare le risorse ospedaliere e garantire migliori risultati per la salute neonatale.

Il progetto si inserisce all'interno di un contesto di crescente attenzione verso la prevenzione delle complicazioni neonatali. La possibilità di prevedere il peso alla nascita dei neonati rappresenta un'opportunità fondamentale per migliorare la pianificazione clinica e ridurre i rischi associati a nascite problematiche, come parti prematuri o neonati con basso peso. Di seguito, i principali benefici che questo progetto porterà all'azienda e al settore sanitario:

### 1. Miglioramento delle previsioni cliniche:

- Il peso del neonato è un indicatore chiave della sua salute. Avere un modello predittivo accurato consente al personale medico di intervenire tempestivamente in caso di anomalie, riducendo le complicazioni perinatali come le difficoltà respiratorie o l'ipoglicemia.

### 2. Ottimizzazione delle risorse ospedaliere:

- Sapere in anticipo quali neonati potrebbero avere bisogno di cure intensive aiuta a organizzare le risorse umane e tecnologiche degli ospedali in modo efficiente. Questo si traduce in una riduzione dei costi operativi e una migliore pianificazione dell'utilizzo delle unità di terapia intensiva neonatale (TIN).

### 3. Prevenzione e identificazione dei fattori di rischio:

- Il modello potrà evidenziare i fattori che maggiormente influenzano negativamente il peso del neonato (come il fumo materno, gravidanze multiple o età avanzata della madre). Queste informazioni sono preziose per la prevenzione e la gestione personalizzata delle gravidanze, permettendo interventi proattivi in caso di rischio elevato.

### 4. Valutazione delle pratiche ospedaliere:

- Attraverso un'analisi comparativa tra i tre ospedali coinvolti, l'azienda potrà identificare eventuali differenze nei risultati clinici, come una maggiore incidenza di parti cesarei in una determinata struttura. Ciò consente di monitorare la qualità delle pratiche e armonizzare i protocolli tra i diversi centri ospedalieri, migliorando la coerenza delle cure.

### 5. Supporto alla pianificazione strategica:

- L'analisi dei dati e le previsioni possono essere utilizzate per prendere decisioni informate non solo a livello clinico ma anche strategico. L'azienda potrà sfruttare queste informazioni per implementare nuove politiche di salute pubblica, garantendo un impatto positivo sui tassi di mortalità e morbidità neonatale.

## Dettagli del Progetto

### 1. Raccolta dei Dati e Struttura del Dataset

Per costruire il modello predittivo, abbiamo raccolto dati su 2500 neonati provenienti da tre ospedali. Le variabili raccolte includono:

- Età della madre:** Misura dell'età in anni.
- Numero di gravidanze:** Quante gravidanze ha avuto la madre.
- Fumo materno:** Un indicatore binario (0=non fumatrice, 1=fumatrice).
- Durata della gravidanza:** Numero di settimane di gestazione.
- Peso del neonato:** Peso alla nascita in grammi.
- Lunghezza e diametro del cranio:** Lunghezza del neonato e diametro craniale, misurabili anche durante la gravidanza tramite ecografie.
- Tipo di parto:** Naturale o cesareo.
- Ospedale di nascita:** Ospedale 1, 2 o 3.
- Sesso del neonato:** Maschio (M) o femmina (F).

L'obiettivo principale è identificare quali di queste variabili sono più predittive del peso alla nascita, con un focus particolare sull'impatto del fumo materno e delle settimane di gestazione, che potrebbero indicare nascite premature.

### 2. Analisi e Modellizzazione

#### Analisi Preliminare

Nella prima fase, esploreremo le variabili attraverso un'analisi descrittiva per comprenderne la distribuzione e identificare eventuali outlier o anomalie.

Inoltre si suggeriranno le seguenti ipotesi con i test adatti:

- in alcuni ospedali si fanno più parti cesarei
- La media del peso e della lunghezza di questo campione di neonati sono significativamente uguali a quelle della popolazione
- Le misure antropometriche sono significativamente diverse tra i due sessi

#### Creazione del Modello di Regressione

Verrà sviluppato un modello di regressione lineare multipla che includa tutte le variabili rilevanti. In questo modo, potremo quantificare l'impatto di ciascuna variabile indipendente sul peso del neonato ed eventuali interazioni. Ad esempio, ci aspettiamo che una maggiore durata della gestazione aumenterebbe in media il peso del neonato.

#### Selezione del Modello Ottimale

Attraverso tecniche di selezione del modello, come la minimizzazione del criterio di informazione di Akaike (AIC) o di Bayes (BIC), selezioneremo il modello più parsimonioso, eliminando le variabili non significative. Verranno considerati anche modelli con interazioni tra le variabili e possibili effetti non lineari.

#### Analisi della Qualità del Modello

Una volta ottenuto il modello finale, valuteremo la sua capacità predittiva utilizzando metriche come  $R^2$  e il Root Mean Squared Error (RMSE). Un'attenzione particolare sarà rivolta all'analisi dei residui e alla presenza di valori influenti, che potrebbero distorcere le previsioni, indagando su di essi.

### 3. Previsioni e Risultati

Una volta validato il modello, lo useremo per fare previsioni pratiche. Ad esempio, potremo stimare il peso di una neonata considerando una madre alla terza gravidanza che partorisce alla 39esima settimana.

### 4. Visualizzazioni

Infine, utilizzeremo grafici e rappresentazioni visive per comunicare i risultati del modello e mostrare le relazioni più significative tra le variabili. Ad esempio, potremmo visualizzare l'impatto del numero di settimane di gestazione e del fumo sul peso previsto.

## Conclusioni

Il progetto di previsione del peso neonatale è un'iniziativa fondamentale per Neonatal Health Solutions. Attraverso l'uso di dati clinici dettagliati e strumenti di analisi statistica avanzati, possiamo contribuire a migliorare la qualità della cura prenatale, ridurre i rischi per i neonati e ottimizzare l'efficienza delle risorse ospedaliere. Questo progetto rappresenta un punto di svolta per l'azienda, consentendo non solo un miglioramento della pratica clinica ma anche l'implementazione di politiche sanitarie più informate e proattive.

## Dataset

Link al dataset: <https://drive.google.com/file/d/1ChfwftuOSH-WLlto7AvV-sQlksGeTq/view>

**Modalità di consegna:** Link a rpubs