MIND-BLOOM

Machine Learning for Prediction and Analysis of Postpartum Depression in Bangladeshi Mothers

Mohammad Ismum | Fatema Hossain
Abrar Mohammed Tanzim Alam | Salma Hossain

PROBLEM STATEMENT

- Predict the likelihood and severity of Postpartum Depression (PPD) in new mothers based on demographic, relational, and psychological factors.
- Develop an interpretable ML system that assists clinicians in identifying at-risk mothers by explaining why the system flags someone as high-risk.
- Study how depression levels change before, during, and after pregnancy, and identify the strongest predictors of postpartum decline.

RELEVANT STATISTICS

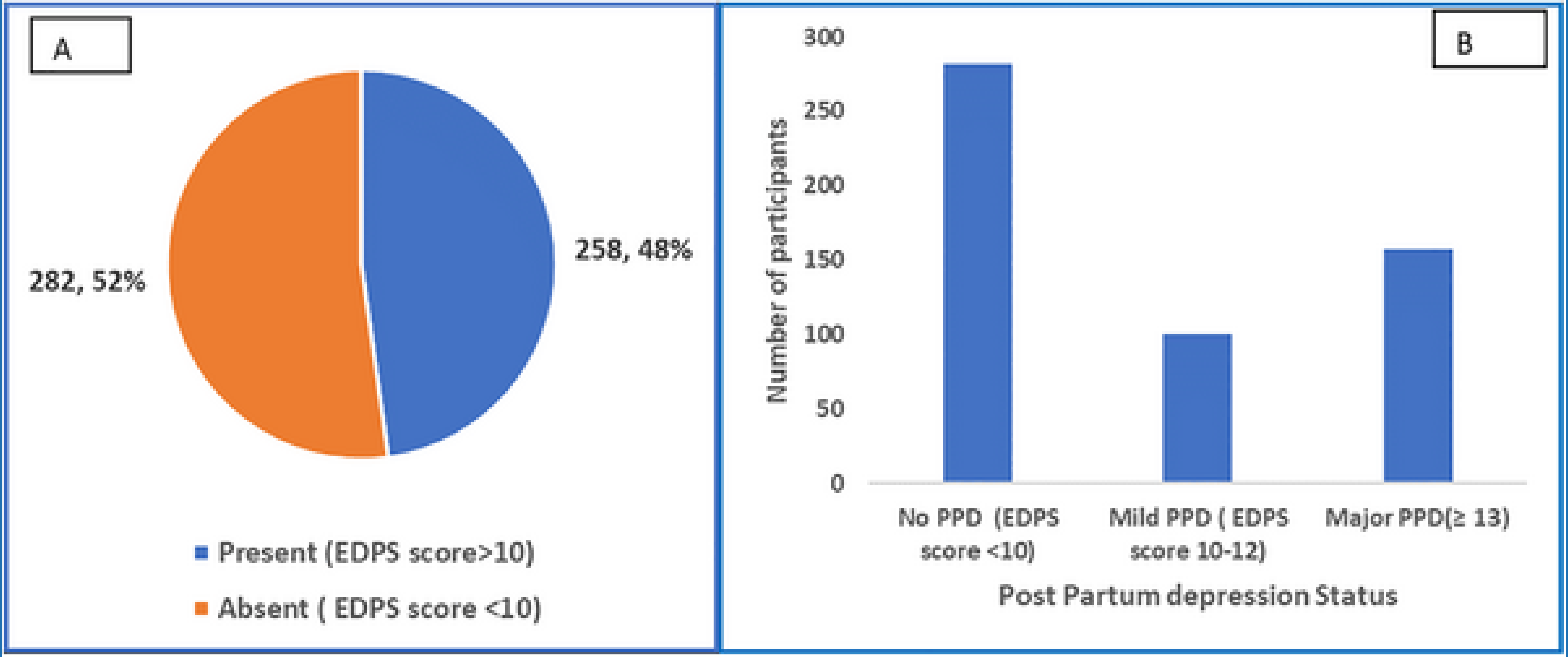
- Around 44% of postpartum women in Bangladesh are at risk of high levels of postpartum depression according to the Edinburgh Postnatal Depression Scale (EPDS). Approximately 28% show moderate to severe depression based on PHQ-9 scores.

- Key risk factors include job loss due to pregnancy, unintended pregnancy, depressive symptoms during pregnancy, intimate partner violence, and a history of miscarriage or child death.

- In rural Bangladesh, the incidence of postpartum depressive symptoms was found to be 8% at 2-3 months postpartum and increased to 18.4% by 6-8 months postpartum.

- Other risk factors specific to Bangladesh include poor socioeconomic status, physical partner violence during pregnancy, anxiety symptoms during pregnancy, and a previous history of depressive symptoms.
- Cultural factors, such as strong male child preference, also influence postpartum depression rates among rural Bangladeshi mothers.

DATA REPRESENTATION



EXPECTED OUTCOMES

- A predictive ML model (Random Forest, XGBoost, or Logistic Regression).
- SHAP/LIME analysis showing top contributors to PPD.
- An interpretable dashboard or web app for screening risk.
- A transparent model for use in clinical and public health settings.
- A visual dashboard summarizing important risk factors.
- Insights into the most critical turning points for mental health intervention.
- Statistical report or visualization of depression evolution patterns.

DATASET DETAILS

Columns (Features)

- Demographics: Age, Residence, Education Level, Marital Status, Occupation
- Economic factors: Monthly income (before/after), Husband's income
- Family & Relationship: Relationship with husband, in-laws, newborn, family type
- Pregnancy history: Number of pregnancies, losses, pregnancy length, mode of delivery
- Health factors: Diseases before/during pregnancy, sleep, anger, breastfeeding, etc.
- Psychological scales: PHQ2 (before/during), PHQ9, EPDS

Label distribution

- PHQ9 Result

Moderate:237(29.6%)
Mild:231(28.9%)
Moderately Severe:132 (16.5%)
Minimal:103(12.9%)
Severe:89(11.1%)
Normal:8{1.0%)

- EPDS Result

High:350()43.8%
Low:260(32.5%)
Medium:190(23.8%)

Data quality

- Verify no missing values (processed file appears complete).

Rows

- 800

DATASET SAMPLE

Age	Residence	Education Level	Marital status	Occupation before latest pregnancy	Monthly income before latest pregnancy	Occupation After Your Latest Childbirth	Current monthly income	Husband's education level	Husband’s monthly income
24	City	University	Married	Student	None	Student	None	University	More than 30000

Addiction	Total children	Disease before pregnancy	History of pregnancy loss	Family type	Number of household members	Relationship with the in-laws	Relationship with husband	Relationship with the newborn	Relationship between father and newborn
None	One	None	None	Nuclear	6 to 8	Neutral	Good	Good	Good

Feeling about motherhood	Recieved Support	Need for Support	Major changes or losses during pregnancy	Abuse	Trust and share feelings	Number of the latest pregnancy	Pregnancy length	Pregnancy plan	Regular checkups
Neutral	High	Medium	Yes	Yes	Yes	1	10 months	No	Yes

Fear of pregnancy	Diseases during pregnancy	Age of newborn	Age of immediate older children	Mode of delivery	Gender of newborn	Birth compliancy	Breastfeed	Newborn illness	Worry about newborn
Yes	None	6 months to 1 year	None	Normal Delivery	Boy	No	Yes	No	Yes

Relax/sleep when newborn is tended	Relax/sleep when the newborn is asleep	Angry after latest child birth	Feeling for regular activities	Depression before pregnancy (PHQ2)	Depression during pregnancy (PHQ2)	PHQ9 Score	PHQ9 Result	EPDS Score	EPDS Result
Yes	Yes	No	Worried	Negative	Negative	14	Moderate	13	High

Feature explanation

Column	Type	Column	Type	Column	Type	Column	Type	Column	Type
Age	Numeric	Occupation After Your Latest Childbirth	Categorical	Relationship with in-laws	Ordinal (Good/Average/Bad)	Diseases during pregnancy	Categorical	Major changes or losses during pregnancy	Boolean
Residence	Categorical	Current monthly income	Ordinal (Range)	Relationship with husband	Ordinal	History of pregnancy loss	Boolean	Feeling about motherhood	Ordinal
Education Level	Categorical	Husband's education level	Categorical	Relationship with newborn	Ordinal	Pregnancy length	Numeric	Recieved Support	Boolean
Marital status	Categorical	Husband's monthly income	Ordinal (Range)	Relationship between father and newborn	Ordinal	Mode of delivery	Categorical	Need for Support	Boolean
Occupation before latest pregnancy	Categorical	Family type	Categorical	Trust and share feelings	Boolean	Birth compliancy	Boolean	Fear of pregnancy	Boolean/Ordinal
Monthly income before latest pregnancy	Ordinal (Range)	Number of household members	Numeric	Disease before pregnancy	Categorical	Pregnancy plan	Boolean	Addiction	Boolean

Feature explanation Cont.

Column	Type	Column	Type	Column	Type
Abuse	Boolean	Worry about newborn	Ordinal	PHQ9 Score	Numeric
Angry after latest childbirth	Boolean	Relax/sleep when newborn is tended	Boolean	PHQ9 Result	Categorical
Feeling for regular activities	Ordinal	Relax/sleep when newborn is asleep	Boolean	EPDS Score	Numeric
Age of newborn	Numeric	Breastfeed	Boolean	EPDS Result	**Target Label**
Gender of newborn	Categorical	Depression before pregnancy (PHQ2)	Categorical (Positive/Negative)		
Newborn illness	Boolean	Depression during pregnancy (PHQ2)	Categorical		

ML PIPELINE:

DEFINING THE PROBLEM

DATA PREPARATION

DATA SPLIT

MODEL SELECTION

MODEL TRAINING

EVALUATION

DEPLOYMENT

DOCUMENTATION

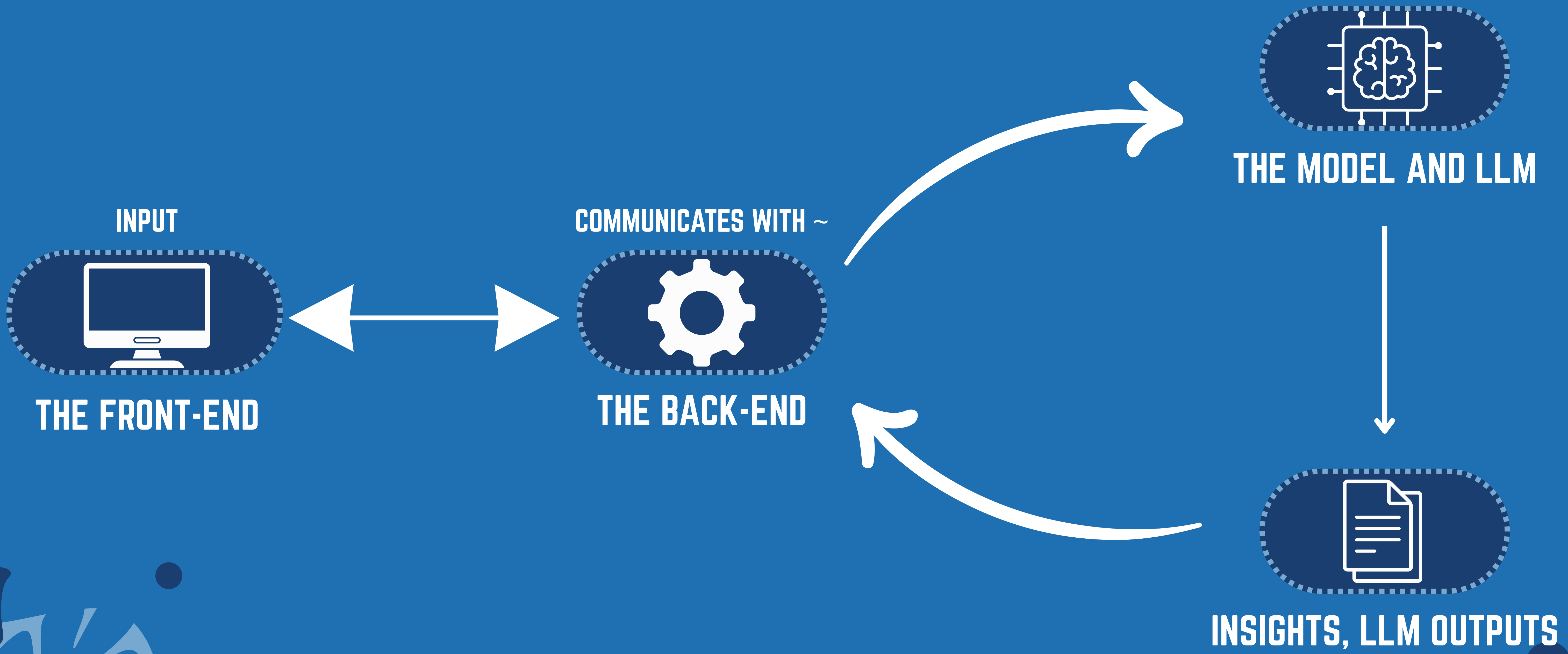
PROBLEM TYPE DEFINITION

- Task type: Supervised Machine Learning — Multiclass Classification

- Input:
  Age, Education, Residence
  Income (before & after childbirth)
  Marital and Family Relationships
  Health Conditions & Pregnancy History
  Emotional and Behavioral Factors
  Previous Depression Indicators (PHQ2)

- Target EPDS Result → Postpartum Depression Severity
  (Categories: Low, Medium, High)

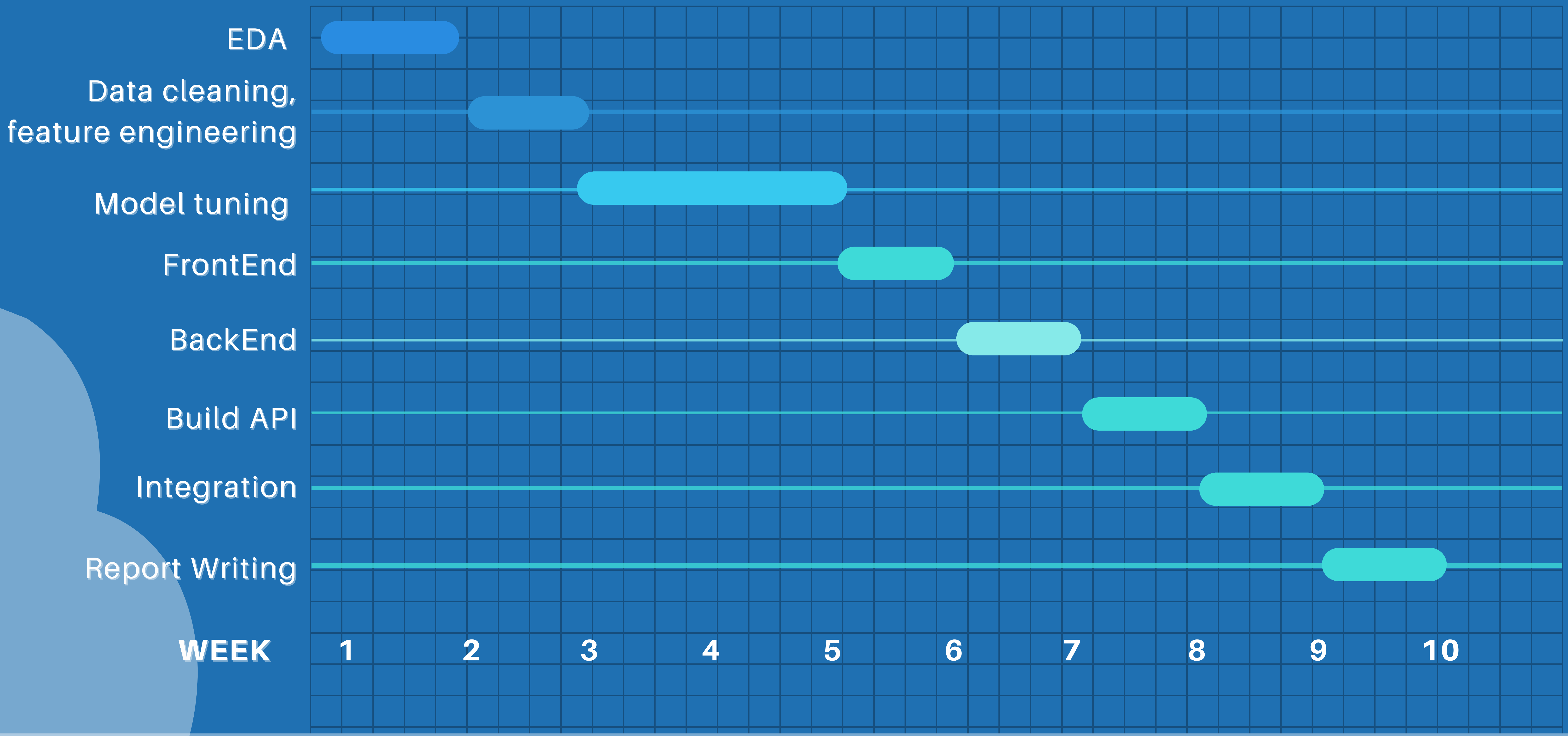
HIGH-LEVEL VIEW OF THE SOFTWARE:



TECHNOLOGY STACK

Data & ML	Python (pandas, scikit-learn, XGBoost/LightGBM), SHAP, NumPy, SciPy
Visualization	matplotlib, seaborn
App API	FastAPI or Flask
Frontend	Flatter, Dart
Database / Storage:	PostgreSQL
Experiment tracking:	MLflow
Version control:	Git + GitHub
Reporting	Jupyter Notebook

GANTT CHART


EDA
Data cleaning,
feature engineering
Model tuning
FrontEnd
BackEnd
Build API
Integration
Report Writing
WEEK
1
2
3
4
5
6
7
8
9
10

WORK DISTRIBUTION AMONGST MEMBERS

Name	ID	Work Distribution
Mohammad Ismum	2212185642	Hyperparameter Tuning - Integrating Model In Backend
Abrar Mohammed Tanzim Alam	2222864042	Pre Tuned Transformer Models - Hybrid Approach -Backend & Database
Fatema Hossain	2222942042	Evaluation Using Traditional ML Techniques - Developing Backend
Salma Hossain	2222943042	Exploratory Data Analysis (EDA) - Designing Frontend

REFERENCE PAPERS

- Prediction of postpartum depression in women: development and validation of multiple machine learning models” (2025)
  https://pubmed.ncbi.nlm.nih.gov/40055720/

- Interpretable Machine Learning Model for Predicting Postpartum Depression: Retrospective Study” (2025)
  https://medinform.jmir.org/2025/1/e58649

- Postpartum depression risk prediction using explainable machine learning algorithms
  https://pubmed.ncbi.nlm.nih.gov/40852354/

PROPOSED NOVELTY

- Use both clinical scales (PHQ/EPDS) and non-clinical socio-psychological indicators (e.g., relationship with husband, sleep quality, financial change) to improve accuracy. Integrate a feature importance analysis to highlight the most influential risk factors — useful for public health interventions.

- Use explainable AI (SHAP, LIME, Decision Trees visualization). Build an AI fairness check across different socio-economic groups.

- Use the PHQ2 before and PHQ2 during pregnancy features to model progression. Employ time-series or sequential modeling (e.g., recurrent models or Markov transitions).

- The website or app collects user feedback on the correctness of predictions to guide future model refinement, making the model self-improving and interactive.

- Combine the text features with the other categorical features. Where we can use a ColumnTransformer to apply TF-IDF to the text column and OneHotEncoder to the other categorical columns.

THANK YOU