# Project Title: BiasX-Ray

**Hackathon:** GNEC 2025

**Target SDG:** SDG 5 - Gender Equality

**Deadline:** May 10, 2025

**🔍 Project Description**

BiasX-Ray is a real-time bias detection and tone analysis tool designed to promote gender-inclusive communication in professional environments. Users can paste or compose team messages, job descriptions, feedback, or emails, and the system will instantly:

* Detect gender-coded language
* Flag subtle microaggressions
* Analyze overall tone
* Offer inclusive rewrite suggestions
* Provide a Bias Score (0-100)

The tool encourages users to reflect on how their communication may be perceived and helps them craft more inclusive and equitable messages.

**💪 Core Features**

1. **Live Text Analysis:**
   * Input box for user messages
   * Real-time bias & tone detection
   * Inline highlights of problematic words or phrases
2. **Gender Bias Detection:**
   * Highlights masculine-coded and feminine-coded terms
   * Suggestions for more neutral alternatives
3. **Microaggression Detector:**
   * Detects common subtle phrases with problematic implications
   * Explanations of why the phrase might be harmful
4. **Tone Analysis:**
   * Classifies text as Friendly, Formal, Aggressive, Anxious, Passive
   * Based on a trained TF-IDF + Logistic Regression model (reuse from Mood2Mail)
5. **Bias Score + Rewrites:**
   * Bias Score (0-100) with feedback on how to improve
   * Rewriting tool for generating bias-free versions
6. **Perspective Shift Feature:**
   * Brief simulated feedback on how the message may feel from a recipient’s point of view

**👨‍💻 Tech Stack**

* **Frontend:** HTML, CSS, JavaScript (Vanilla)
* **Backend:** Python, Flask
* **ML:** Scikit-learn (TF-IDF + Logistic Regression), optional pattern-based rules
* **Data:**
  + Gender-coded word list
  + Microaggression phrase database
  + Manually labeled sample text for tone classification

**🔢 Modules Breakdown**

**1. Bias Detection Engine**

* **Input:** User message (string)
* **Output:**
  + List of gendered words with type (masculine/feminine)
  + List of microaggression flags with explanation

**2. Tone Classifier**

* **Input:** User message
* **Output:** One of 5 tones
* **Model:** TF-IDF + Logistic Regression trained on labelled dataset (275 samples, scalable)

**3. Flask API**

* Endpoints:
  + /analyze: accepts text, returns bias highlights, tone, score, rewrites

**4. Frontend UI**

* Textarea with real-time analysis
* Highlighted text elements
* Bias meter (score visual)
* Suggestions panel with corrections
* Optional: Perspective switcher dropdown

**⏳ Timeline**

| **Dates** | **Milestones** |
| --- | --- |
| Apr 18–20 | Build bias & microaggression lexicon |
| Apr 21–24 | Tone detection model + testing |
| Apr 25–27 | Flask API for /analyze endpoint |
| Apr 28–30 | Frontend: input box, highlights, bias UI |
| May 1–4 | Suggestions + rewrite functionality |
| May 5–8 | UI polish, mobile view, user testing |
| May 9–10 | Final deployment & hackathon submission |

**🚀 Launch Plan**

* Deploy on **Render** or **Railway**
* Add **demo video + GitHub repo**
* Prepare a **one-liner pitch** and **problem-solution slide** for judging

Let me know whenever you want to start building the lexicon or the tone model — I’ll guide you step by step.

For tone model --- use mood2mail's model