Bank of Baroda Hackathon

Theme: Hybrid Identity Monitoring & Deepfake-Resistant Verification

Team: GaurNitai

Bhavya Garg (220296) & Saagar K V (220927)

1. Problem Statement & Relevance to Banking

With the rise of AI-driven fraud, particularly deepfakes and synthetic identities, financial institutions face a serious challenge in ensuring secure and trusted digital onboarding. Video KYC, a key element of customer acquisition in modern banking, is highly vulnerable to face swaps, voice clones, and lip-sync manipulation.

For a bank, this risk is not only about fraud losses but also about customer trust, regulatory compliance, and reputation. We identified this gap as highly relevant to the future of secure banking.

2. Our Solution: TrustNet 360^{o}

We proposed a **multi-layered verification system** designed to resist deepfake attempts and continuously assure identity throughout the user's banking journey. The core components are:

1. **Dynamic Biometric Interrogation (DBI):**
   Randomized prompts like *"Say 72-15-93 while looking top-right"*. This allows us to simultaneously check **facial expression, lip sync, and attention cues**.

   The inspiration came partly from **CAPTCHA and reCAPTCHA systems**, where unpredictability makes it hard for bots to automate responses. Similarly, DBI makes it significantly harder for deepfakes to generate correct responses in real-time.

   We also drew parallels with **Dynamic OTP systems**, where the challenge is unique for each user/session, ensuring stronger security.

2. **Ensemble Detection Model:** Combines facial landmark tracking, lip-audio coherence, and environmental consistency checks to flag manipulation attempts.

3. **Trust Score Engine:** Each interaction is scored based on authenticity signals. Depending on the score, the system can allow smooth flow, trigger re-authentication, or escalate to manual review.

4. **Conceptual Explorations:**

   - ***Uncanny Valley as Prediction Error****:* Inspired by cognitive science research, we explored the idea that deepfakes fail to replicate authentic human motion. By modelling prediction errors in facial movement, fakes can be revealed.

   - ***Affective Dissonance****:* Real emotions leak through micro-expressions and subtle muscle activations (e.g., Duchenne smile). Deepfakes usually miss these cues. Detecting such incoherence provides a new dimension of security.

   - **Remote Photoplethysmography (rPPG):** Uses a standard camera to estimate physiological signals like heart rate and oxygen from subtle skin color changes, adding a non-contact biometric factor. While accuracy can be affected by lighting and motion, it adds another layer of liveness assurance.

3. Our Journey Through the Hackathon

When we first came across the hackathon, it coincided with our mid-semester exams, which left us with very limited time. Initially, we could only attempt some simple prototypes and brainstorming. Fortunately, when the deadline was extended and our exams ended, we shifted gears into intensive research.

We studied the existing challenges in banking KYC, the rules and regulations of RBI, and surveyed current industry solutions. This gave us a realistic picture of the gaps that still exist.

Coming from a biological sciences background, one of us drew inspiration from human-specific behaviours and neurological mechanisms — such as micro-expressions and prediction errors — which provided novel angles for deepfake detection beyond pure computer vision.

We also had to navigate the fact that cybersecurity was a new domain for both of us. A significant part of our journey was understanding the key terminologies and frameworks of the field before we could design our system.

Our implementation was a prototype-level exploration rather than a fully production-ready system. While we attempted to create a basic interface, our main focus was on laying down the core foundation of our proposed solution. Overall, it was a journey of rapidly building domain understanding, bridging interdisciplinary knowledge, and prioritizing what mattered most in the time we had.

4. Role of Bank of Baroda, IIT Kanpur & SIIC

- **Bank of Baroda** provided the real-world banking context and inspired us to think about deployment in regulated, high-trust environments.
- **IIT Kanpur and SIIC:** offered valuable technical feedback on feasibility and the science behind detection models. SIIC guided us through the registration and proceedings of the competition.

5. Outcomes, Learnings & Future Scope

Key Learnings:

- Deepfake detection requires going beyond artifact spotting — integrating human behavioural science with machine learning opens powerful new directions.
- Adaptive trust scoring can strike the right balance between user security and seamless banking experience.
- Entering a new domain like cybersecurity requires upfront investment in learning terminology, frameworks, and real-world fraud cases.

Outcomes:

- A conceptual framework that incorporates prediction error (uncanny valley) and affective dissonance (emotional leakage).
- Interdisciplinary insights by applying biological and neurological principles to digital identity verification.

Future Plans:

- Expand our prototype into a modular API service for banks and fintechs.
- Validate our innovative detection approaches (uncanny valley, emotional coherence) on larger datasets and controlled experiments.
- Collaborate with banking partners for pilot projects and strengthen technical depth with more time and resources.

6. Closing Reflection

This Hackathon gave us a unique chance to work at the intersection of AI, security, and banking. Our solution is not the final word, but a starting point toward resilient, human-centric digital identity verification.

We leave with both a prototype and a roadmap, and more importantly, with the confidence that interdisciplinary thinking—technical, behavioural, and regulatory—will shape the future of secure banking.