**PROJECT PROBLEM STATEMENT:**

The escalating prevalence of social media platforms as communication channels has provided an avenue for certain individuals to exploit these platforms by disseminating phishing links and propagating hate speech within comments, particularly on widely used platforms such as Instagram. Presently, there has been a notable surge in user engagement on Instagram. Consequently, ensuring user privacy, safety, and content quality has emerged as a pressing concern. Design and develop the machine learning model that can analyze the Instagram comments and flag them with two categories as below.

**Phishing comments:**

Detects comments that contain phishing links or fraudulent content.

Examples: “Click this link for free followers,” “Win a brand-new iPhone here,” etc.

**Vulgar or hate speech:**

Identify comments with inappropriate language, hate speech, or offensive content.

Examples: Profanity, racial slurs, threats, etc.