Team Name: **HexaSentinel**

Members:
- Nabhonil Bhattacharjee(Student)
- Jeet Biswas (Student)
- Prof. Palash Dutta(Professor)

Problem Statement: **Supervised Learning for City-Level IP Geolocation**

TABLE OF CONTENTS

Blog link

Introduction

- **Theme:** Implementation and Testing of Selected Internet-Drafts / RFCs
- **Focus Areas:** Supervised IP Geolocation & Measurement Calibration
- **Organized by:** Advanced Internet Operations Research in India (AIORI)
- **Collaborating Institutions:** Heritage Institute of Technology, Kolkata
- **Date:**11/2025
- **Prepared by:**

Name	Designation	Institution
Nabhonil Bhattacharjee	Student	Heritage Institute of Technology, Kolkata
Jeet Biswas	Student	Heritage Institute of Technology, Kolkata
Prof. Palash Dutta	Professor	Heritage Institute of Technology, Kolkata

Contact: nabhonilbhattacharjee@gmail.com

- ## Executive Summary

Team HexaSentinel developed GeoNex, a calibrated, ML-driven city-level IP geolocation system implementing IETF RFC 8805, 2330, 2681, 7679, 7680, 792, 4443, 4271, 6793, 9081–9083, 1035, 3152, and 3596. The project integrates supervised learning, active probing, and MLOps for verifiable location prediction with uncertainty quantification. GeoNex achieved 70–80% real-world accuracy with confidence radius estimation, enabling trustworthy Internet infrastructure analytics.

- ## Overview

This project operationalizes RFC 8805 and IPPM standards to advance active delay and loss measurement. We are developing reproducible ML pipelines that integrate network metrics with geospatial data, contributing calibration modules and test datasets directly to open-source repositories.

Beyond technical implementation, this initiative generates critical feedback for the IETF IPPM and MAPRG working groups, bridging the gap between standards and deployment. Ultimately, this work solidifies operational expertise in Internet measurement standards within the AIORI ecosystem.

- **Objectives**
  - Implement RFC 8805 (self-published IP geolocation) and IPPM RFCs for active delay/loss measurement.
  - Develop reproducible ML pipelines integrating network measurement and geospatial data.
  - Contribute calibration modules and test data to open-source repositories.
  - Generate implementation feedback to IETF IPPM and MAPRG working groups.
  - Build operational expertise in Internet measurement standards within the AIORI ecosystem.

- **Scope and Focus Areas**

Focus Area	Relevant RFCs / Drafts	Open Source Reference
IP Geolocation (City-level)	RFC 8805, RFC 2330, RFC 2681, RFC 7679, RFC 7680, RFC 9081–9083	GeoLite2, RIPE Atlas API, LightGBM, FastAPI
Active Probing	RFC 792, RFC 4443	Go-based ICMP Probe Manager
BGP & RDAP Integration	RFC 4271, RFC 6793, RFC 9081	pybgpstream, RDAP Python Library
Reverse DNS & DNS Features	RFC 1035, RFC 3152, RFC 3596	dnspython

- **Sprint Methodology**

A four-phase sprint workflow was followed using virtual testbed and Dockerized environments:
  - RFC / Draft Selection – RFCs related to IP measurement and geolocation chosen.
  - Sprint Preparation – Dataset creation from GeoLite2, RIPE Atlas, National Internet Backbone, and RDAP sources.
  - Implementation Phase – ML model training and calibration with active network data.
  - Interoperability Testing – Cross-validated with multiple AIORI nodes for accuracy.
  - Documentation & Contribution – Open-sourced pipeline via GitHub.
  - Post-Sprint Reporting – Results and performance metrics shared with AIORI mentors.

- **Activities and Implementation**

Date	Activity	Description
05/10/2025	Sprint 1: Data Acquisition	Integrated IP→city datasets, ASN, and RTT data
12/10/2025	Sprint 2: ML Calibration	Implemented isotonic regression for confidence radius calibration
20/10/2025	Sprint 3: Active Probe Manager	Built Go-based RTT collection using ICMP (RFC 792/4443)
27/10/2025	Sprint 4: API & Visualization	Deployed FastAPI inference with Tkinter frontend

- ## Results and Findings

GeoNex achieved Top-1 city accuracy of 77% and median geo-error of 35 km on held-out test data.

  - Confidence radius covered 92% of ground truth within predicted bounds.
  - RTT-based feature integration improved prediction stability across ASNs.

- ## Open Source Contributions

Project	Contribution	Status	Link
GeoNex Repository	MLOps pipeline & calibration code	Merged	https://github.com/nabhocharger69/HexaSentinel

- ## Collaboration with IETF WGs

Feedback and implementation notes were shared with IETF MAPRG and IPPM WGs, highlighting calibration use cases under RFC 2330. Proposed a draft idea on "Confidence Radius Metrics for IP Geolocation Models."

- ## Impact and Future Work

GeoNex's outcomes will integrate with AIORI-IMN for measurement visualization and anomaly detection. Future work includes:

  - Extending calibration for IPv6 datasets and Anycast detection.
  - Publishing evaluation results as an IETF Internet-Draft (AIORI contribution).
  - Incorporating privacy safeguards and adaptive regional learning.

- # Technical Blog Series & Dev Diary

- ## Lead Paragraph

In the AIORI-2 Hackathon, Team HexaSentinel tackled the challenge of city-level IP geolocation—transforming Internet addresses into calibrated, trustworthy coordinates using RFC 8805 and IPPM measurement standards. The result: GeoNex, a supervised learning system that quantifies its own uncertainty.

- ## Background and Motivation

Traditional IP databases are static and inaccurate, offering no reliability measure. RFC 8805 and IPPM frameworks (RFC 2330, 2681) provide mechanisms for collecting verifiable geolocation data. GeoNex implements these in an ML workflow to create auditable, confidence-aware predictions—critical for routing optimization, fraud prevention, and network research.

- ## Technical Implementation

1. Setup and Tools

  - Node: Heritage Institute of Technology, Kolkata
  - OS: Ubuntu 24.04 LTS
  - Languages: Python, Go, C
  - Libraries: LightGBM, FastAPI, DVC, pybgpstream, dnspython
  - Measurement Tools: RIPE Atlas API, ICMPv4/v6 Probes, Traceroute, Wireshark

2. Implementation Steps

  - Collected GeoLite2, ASN, RDAP, and RTT data into GeoJSON.
  - Trained gradient-boosted models (LightGBM) to predict city and confidence radius.
  - Calibrated probabilities using isotonic regression.
  - Exposed real-time inference via FastAPI and visualized results in Tkinter.

3. Challenges Faced

  - RTT normalization across diverse network vantage points.
  - Lack of ground truth for smaller cities.
  - Integrating multi-language modules (Go, C, Python) under one MLOps pipeline.
  - Lack of recent data and insufficient Indian data.

- ### Results and Observations

Test	Metric	Observation	Note
City Accuracy	77%	Matches target dataset distribution	Improved over baseline (63%)
Geo Error (Median)	35 km	Within predicted confidence radius	Calibrated via isotonic regression
RTT Stability	±2 ms variance	Stable across AIORI nodes	Indicates reliable measurement

Command Example (Just a visual):

```
$ curl -X GET 'http://geonex.aiori.in/predict?ip=103.97.x.x'
{"city":"Kolkata","confidence":0.81,"radius_km":32.4}
```

- ### Lessons Learned
  - Implementing RFC-based measurement improves trust in ML predictions.
  - Combining active and passive data increases interpretability.
  - Real-world IETF collaboration bridges research and deployment.

Reporting and Standards Mapping

Team Name	Institution	Project Title	Focus Area
HexaSentinel	Heritage Institute of Technology, Kolkata	GeoNex: Calibrated City-Level IP Geolocation	✅ Other (Measurement / IP

Date: November 2025

1. Standards Reference

RFC / Draft No.	Title / Area	Lifecycle Stage	How This Work Relates
RFC 8805	Self-Published IP Geolocation Data	Internet Standard	Implemented full RFC pipeline for IP→city data
RFC 2330 / 2681 / 7679 / 7680	IP Performance Metrics Framework	Proposed/Standard	Used RTT, delay, and loss metrics in ML features
RFC 792 / 4443	ICMP/ICMPv6	Internet Standard	Implemented active probes to collect RTT samples
RFC 9081–9083	RDAP Protocol	Proposed Standard	Integrated ASN and prefix ownership data
RFC 4271 / 6793	BGP Routing	Internet Standard	Extracted origin ASN context for generalization

2. Impact on Standards Development

Question	Response
Does this work support, extend, or validate an existing RFC?	Yes, validates RFC 8805 by integrating active IPPM and RDAP data into ML-driven geolocation.
Could it influence a new Internet-Draft or update sections of an RFC?	The confidence radius metric could extend RFC 8805 for ML-based uncertainty calibration.
Any feedback or data shared with IETF WG mailing lists?	Shared observations with MAPRG and IPPM lists during AIORI-2 discussions.
Planned next step	Prepare Internet-Draft proposal: *"Confidence Metrics for City-Level IP Geolocation."*

- **Acknowledgments**

Special thanks to AIORI mentors, Heritage Institute of Technology, and RIPE Atlas community for data access and guidance. Appreciation to the AIORI-2 organizing team for fostering open Internet research collaboration.

- **References**
  - RFC 8805 – Self-Published IP Geolocation Data
  - RFC 2330, 2681, 7679, 7680 – IP Performance Metrics Framework
  - RFC 9081–9083 – RDAP Query Protocols

- **Reflections from the Team**
  - **Nabhonil (Lead):** "Working with real Internet data made theory meet engineering."
  - **Jeet:** "Isotonic calibration was the turning point for trustworthy AI."
  - **Mentor:** Prof. Palash Dutta "Data reproducibility isn't optional—it's Internet infrastructure hygiene."