PRODUCT REQUIREMENTS DOCUMENT

# Navi Mumbai House Price Predictor

*AI-powered residential real-estate valuation platform for the Navi Mumbai market*

| | |
|---|---|
| **Document Version** | 1.0 |
| **Status** | Draft |
| **Date** | February 2026 |
| **Owner** | Product Management |
| **Stakeholders** | Engineering, Data Science, Design, Sales & Marketing |

## 1. Executive Summary

The Navi Mumbai House Price Predictor is a machine-learning-powered web and mobile application that provides accurate, real-time property valuations for residential units across Navi Mumbai's key micro-markets — including Vashi, Kharghar, Belapur, Panvel, Airoli, Ghansoli, Nerul, and Seawoods. The platform targets homebuyers, sellers, real-estate agents, and financial institutions, enabling data-driven decisions in one of India's fastest-growing urban corridors.

The product will ingest structured property data, geospatial signals, and macroeconomic indicators, then surface instant price estimates via a clean consumer-facing interface and a REST API for B2B integrations.

## 2. Problem Statement

### 2.1 Market Context

Navi Mumbai's residential market spans diverse micro-markets with wide price variation — a 2 BHK in Kharghar may differ by 30–50% from a comparable unit in Panvel or Airoli. Buyers and sellers routinely rely on broker word-of-mouth, leading to:

- Information asymmetry between buyers and professional brokers
- Overpriced listings that languish on the market for months
- Underpriced distressed sales that harm sellers
- Mortgage under/over-valuation risk for banks and HFCs

### 2.2 Existing Gaps

National portals (MagicBricks, 99acres, Housing.com) provide indicative price bands but lack Navi Mumbai micro-market granularity, recency, and explainability. No dedicated tool currently combines hyperlocal data (CIDCO sector, node, floor, amenity quality) with real-time market signals.

## 3. Goals & Success Metrics

| Goal | Success Metric (12-month target) |
|---|---|
| Prediction accuracy | Median Absolute Percentage Error (MdAPE) ≤ 8% vs. registered transaction values |
| User adoption | 50,000 monthly active users by month 12 |
| API adoption | 20 B2B integrations (banks, proptech platforms) |
| Valuation volume | 5 lakh valuations served per month |
| Data freshness | Price model retrained at least weekly |
| Page speed | Prediction result returned in < 2 seconds (p95) |

## 4. Scope & Assumptions

### 4.1 In Scope — V1

- Residential properties only: apartments/flats, independent floors
- Geography: 12 nodes of Navi Mumbai (CIDCO planning area)
- Property types: 1 RK, 1 BHK, 2 BHK, 3 BHK, 4+ BHK
- Price prediction via web form and REST API
- Comparable sales display (last 10 registered transactions in locality)
- Trend charts: 3-month, 6-month, 1-year price index per node
- Basic report export (PDF)

### 4.2 Out of Scope — V1

- Commercial, industrial, or agricultural properties
- Areas outside CIDCO-planned Navi Mumbai (e.g., Panvel Municipal Corporation fringe)
- Rental price prediction (deferred to V2)
- Property listing marketplace
- Automated Valuation Model (AVM) certificate for legal/loan use without human review

## 5. User Personas

| Persona | Homebuyer / Seller | Real-Estate Agent | Bank / HFC Loan Officer |
|---|---|---|---|
| **Goal** | Know fair market value before negotiating | Price listings accurately; win more mandates | Validate collateral value for home loans |
| **Pain Point** | Broker quotes vary wildly; no trusted reference | Manually tracks WhatsApp/MagicBricks data | Physical inspection is slow and expensive |

# 6. Feature Requirements

## 6.1 Instant Price Estimator (Core)

Users enter property details via a guided form. The system returns a predicted price range (low / mid / high), a confidence interval, and key value drivers.

**Required input fields:**

- Node / locality (dropdown: Vashi, Kharghar, Belapur CBD, Belapur Sector, Nerul, Seawoods, Airoli, Ghansoli, Kopar Khairane, Sanpada, Panvel, Ulwe)
- Society / project name (autocomplete)
- Configuration: 1 RK / 1 BHK / 2 BHK / 3 BHK / 4 BHK+
- Carpet area (sq ft) — RERA-defined carpet area
- Floor number & total floors in building
- Age of property (years)
- Furnishing status: Unfurnished / Semi-furnished / Furnished
- Parking: None / 1 / 2+
- Facing: East / West / North / South / Corner
- Premium amenities (multi-select): Gym, Swimming Pool, Clubhouse, Gated Security, CCTV, Lift

**Output fields:**

- Estimated price range: ₹X lakh – ₹Y lakh (mid-point highlighted)
- Per sq ft rate: ₹A – ₹B/sq ft
- Confidence band: ±Z%
- Top 3 value drivers (SHAP-based feature attribution in plain language)
- Comparable transactions: last 5–10 IGR-registered deals within 500m & 90 days
- Node price trend sparkline (12 months)

## 6.2 Trend & Analytics Dashboard

- Price index chart per node (monthly, indexed to Jan 2022 = 100)
- Heatmap: price per sq ft across Navi Mumbai nodes (Google Maps overlay)
- Supply-demand indicator: listing count vs. absorption rate
- New project launch tracker (RERA-filed projects in last 90 days)

### 6.3 Saved Valuations & Alerts

- Registered users can save up to 20 valuations
- Price alert: notify by email/SMS if estimated price of a saved property changes by ≥5%
- Portfolio view: aggregate estimated value of multiple saved properties

### 6.4 PDF Valuation Report

- One-page report with property details, estimated value, comparables, and market context
- Disclaimer: 'For indicative purposes only; not a certified valuation report'
- Branding & date-stamped

### 6.5 REST API (B2B)

- POST /v1/valuations — submit property details, receive JSON prediction
- GET /v1/market/trends — time-series price index by node
- GET /v1/transactions — recent registered transactions by locality
- OAuth 2.0 / API key authentication
- Rate limits: 1,000 calls/day (standard), 50,000/day (enterprise)
- SLA: 99.5% uptime; < 2 second p95 latency

## 7. Data Sources & ML Model

### 7.1 Data Sources

| Data Source | Content | Update Frequency |
|---|---|---|
| Maharashtra IGR (Inspector General of Registration) | Registered sale deeds: area, price, date, property details | Daily |
| RERA Maharashtra | Project approvals, carpet area declarations, completion status | Weekly |
| CIDCO | Sector plans, amenity locations, development status | Monthly |
| Google Maps / OSM | POI: stations, schools, hospitals, malls, distance features | Monthly |
| Listing portals (scrape) | Active listing prices, days on market, developer launches | Daily |
| RBI / NHB | Repo rate, home loan rates, HPI index | Monthly |

### 7.2 ML Model Architecture

- Primary model: Gradient Boosting Regressor (XGBoost / LightGBM) trained on IGR transaction data
- Features: ~60 engineered features — structural, locational (node, proximity to metro/station), temporal, macroeconomic
- Target: price per sq ft (log-transformed) — model predicts log(₹/sq ft); back-transformed for output
- Validation: time-based train/test split; last 3 months held out for evaluation
- Explainability: SHAP values computed per prediction to surface top value drivers
- Retraining cadence: automated weekly pipeline; alert if MdAPE degrades > 2pp vs. baseline
- Fallback: if training data < 30 transactions in micro-segment, use node-level model with similarity weighting

## 8. Non-Functional Requirements

| Category | Requirement |
|---|---|
| **Performance** | API p95 latency < 2s; web page LCP < 3s on 4G mobile |
| **Scalability** | Support 10,000 concurrent users; horizontally scalable on Kubernetes |
| **Availability** | 99.5% uptime; max planned downtime 4 hrs/month (midnight maintenance window) |
| **Security** | OWASP Top 10 compliance; data encrypted at rest (AES-256) and in transit (TLS 1.3) |
| **Data Privacy** | No PII collected for anonymous valuation; registered users DPDP Act 2023 compliant |
| **Accessibility** | WCAG 2.1 AA; mobile-first responsive design; Hindi language option in V2 |
| **Observability** | Centralized logging (ELK); Prometheus + Grafana dashboards; PagerDuty alerting |

## 9. Release Roadmap

| Phase | Timeline | Deliverables |
|---|---|---|
| **Alpha** | Months 1–2 | Internal prototype; XGBoost model trained on IGR 2022–2025 data; 3 nodes (Vashi, Kharghar, Nerul); web form only |
| **Beta** | Months 3–4 | All 12 nodes; comparables panel; trend charts; closed beta with 50 real-estate agents |
| **V1 Launch** | Month 5 | Public web app; PDF report; registered accounts; price alerts; PR launch |
| **V1.1** | Months 6–7 | REST API GA; B2B onboarding (banks); heatmap; SHAP explainability improvements |

| | | |
|---|---|---|
| **V2** | Months 9–12 | Rental prediction; Hindi UI; mobile app (React Native); portfolio tracker; advanced analytics dashboard |

## 10. Risks & Mitigations

| Risk | Severity | Mitigation |
|---|---|---|
| IGR data access restricted or paywalled | **High** | Engage MahaRERA/NIC for data-sharing MOU; fallback to listing price data with discount model |
| Model accuracy insufficient for broker/bank trust | **High** | Strict MdAPE ≤ 8% launch gate; display confidence intervals; phased B2B rollout with manual review layer |
| Sparse transaction data in newer nodes (Ulwe, Dronagiri) | **Medium** | Spatial interpolation from adjacent nodes; synthetic augmentation with listing-to-transaction ratio model |
| Data scraping blocked by listing portals | **Medium** | Commercial data agreements with portals; rotate IPs / respect robots.txt; internal data contributor incentive program |
| Regulatory: SEBI / RBI classifies AVM as regulated activity | **Low** | Prominent disclaimer; no loan-use certification; legal review before B2B banking launch |

## 11. Open Questions

- Should the V1 model distinguish CIDCO-built vs. private developer properties explicitly, or handle it via feature engineering?
- What is the monetisation model — freemium (limited free valuations), B2B API licensing, or lead-gen for brokers?
- Will we partner with a licensed valuer to offer certified reports (a stronger B2B value prop for banks)?
- How should we handle super-luxury properties (₹5 Cr+) where comparable sales data is thin?
- Do we build mobile apps natively (Android-first given market) or defer to PWA in V1?

## 12. Approvals

| Role | Name | Signature & Date |
|---|---|---|
| Product Manager | | |
| Engineering Lead | | |
| Data Science Lead | | |
| Design Lead | | |
| Legal / Compliance | | |

*Document prepared by Product Management. For questions contact the document owner.*