PRIYANKA PRAKASH NAIR

+1 (646) 295-0408 | ppn2115@columbia.edu | New York, NY | linkedin.com/in/ppn511 | github.com/ppn2115

SUMMARY

Data scientist with a product manager's instinct, with ~5 years at American Express and AI/ML research at Columbia, building models end to end; graph-based entity resolution for ~6M customers in 70+ countries, anomaly detection on highly imbalanced data, an insider trading signal that beats its baseline out of sample over 22 years and a code-verified LLM scheduling app with evaluation guardrails

WORK EXPERIENCE

Memorial Sloan Kettering Cancer Center **New York, NY**
Business Process Analysis and Optimization Intern *Jun 2026 - Aug 2026*

- Helped secure ~$3M in renovation funding and reframed a $5.9M compliance risk by isolating 86% as a data error via root cause analysis and production Power BI/DAX dashboards, in an internship won by diagnosing a class case study's $15.9M revenue gap
- Sequenced $3.7M in pricing actions across 1,062 of 1,467 units, surfacing $86K in quick wins, by building a 4-tier prioritization framework in Excel and aligning Legal, Finance, IT, and Housing Ops on an SVP-approved roadmap

American Express **Gurugram, India**
Senior Associate, Analytics & Digital Product Management *Aug 2023 - Jun 2025*

- Averted multi-million-dollar regulatory exposure across ~4M Eurozone accounts in an enterprise KYC program by defining requirements and data architecture for a customer-level database with relationship roll-up logic
- Reduced under- and over-linkages by 2-3% across 200M+ relationships by building a graph-based entity resolution pipeline (connected-components clustering in PySpark and NetworkX) that links each customer's relationships to one customer ID
- Scaled the Search API into Amex's system of record as product owner, unlocking 10+ integrations in 4 sprints and cutting error rates by 30% by consolidating 15+ legacy pipelines and leading a 15+ member scrum team
- Decreased monthly Operational Risk Events on linkage issues by 20% by partnering with data stewards to launch SQL/Splunk dashboards and alerts across the new card account journey, shifting issue management from reactive to proactive

Associate, Analytics & Digital Product Management *Sep 2020 - Jul 2023*

- Raised record-linkage accuracy from ~99% to ~100% in production APIs by re-engineering a probabilistic model (fuzzy name, DOB, address, phone, government ID features) into a deterministic ruleset, validated against an analytical matching algorithm
- Cut customer-harm incidents by 80% for ~6M customers in 70+ countries by revamping the Global Matching Framework through a phased historical refresh of customer demographics data with credit risk, marketing, compliance, loyalty, and servicing teams

PROJECTS

LionMeet Connect: LLM Scheduling App (Columbia University, Product Management with AI) *Sep 2026 - Oct 2026*

- Launched a live AI scheduling app for Columbia teams of up to 9 with an "AI proposes, code verifies" design: Claude interprets typed or spoken constraints and ranks options, while deterministic code verifies every slot against hard rules
- Gated releases on a 20-case AI evaluation (20/20 pass), 44 API regression tests, and randomized rule tests; shipped user-feedback features and made scheduling computation 27x faster

Scienaptic AI: Anomaly Detection Models (Columbia University Research Collaboration) *Jan 2026 - May 2026*

- Achieved 86.49% out-of-time Recall@1% and >90% recall at a 1 to 2% flag rate on 284,807 transactions (0.17% fraud rate) by benchmarking XGBoost (200-trial Bayesian tuning, SHAP) against one-class SVDD and DBSCAN
- Selected models on recall at the 1% operating threshold rather than AUC, matching how many flags a fraud team can review, and diagnosed kernel bandwidth as SVDD's binding constraint

Morgan Stanley: Insider Trading Signal Extraction (Columbia University Research Collaboration) *Jan 2026 - May 2026*

- Engineered an end-to-end insider-signal pipeline across 2,506 companies and 22 years of SEC Form 4 filings, validated over 268 monthly walk-forward windows, by building seniority-tier classification, a net insider buying ratio, and quintile ranking
- Achieved a 52.1% win rate ($p < 0.05$) and +4.35% 3-month alpha by designing Filing Lag Urgency and Peer Contagion signals and combining them into a z-scored composite

NYC Yellow Taxi Tip Prediction (Columbia University Data Analytics Project) *Oct 2025 - Dec 2025*

- Identified fare, drop-off location (Google Maps API), and scraped weather/time data as the top tip predictors by comparing 5 models across 30M+ NYC TLC trips and 649 engineered features; Lasso with 12 predictors performed best

EDUCATION

Columbia University in the City of New York **New York, NY**
M.S. in Business Analytics, GPA: 3.89/4.00 *Aug 2025 - Dec 2026 (Expected)*
Coursework: Data Analytics & Machine Learning, AI Applications in Finance, Optimization Models, Product Management with AI

St. Stephen's College, University of Delhi **Delhi, India**
B.Sc. (Honors) Mathematics, GPA: 3.92/4.00 (Top 5%) *Jul 2017 - Oct 2020*
Coursework: Linear Algebra, Probability Theory and Statistics, Calculus (Multivariate), Linear Programming and Game Theory

SKILLS

Machine Learning: XGBoost, Random Forest, Regularized Regression, SVDD, DBSCAN, Graph Clustering, A/B Testing
GenAI & Modeling: LLMs, RAG, Structured Outputs, LLM Evaluation, Feature Engineering, SHAP, Backtesting
Product: Roadmapping, Backlog & User Stories, Acceptance Criteria, Agile/Scrum, UAT, KPI Design, Stakeholder Management
Tools: Python (Pandas, scikit-learn, NetworkX), SQL, PySpark, Kafka, REST APIs, Postman, Splunk, Power BI, Jira, Git