# B2B LEAD SCORING ENGINE

By - Likhith Bhargav

# PROBLEM & TECHNICAL APPROACH

**Problem Statement:**
- Business challenge: Too many undifferentiated leads → inefficient sales outreach
- Goal: Automatically score leads by conversion likelihood

**Tech stack**:
- Snowflake (data warehousing & ingestion)
- Python notebook (feature engineering & RF model, PFI explainability)
- Power BI (interactive dashboard)
- n8n (automated alerts & data pipeline)

```python
# Firmographics + lead_age
fg = ['EMPLOYEE_COUNT','ANNUAL_REVENUE','lead_age'] + [c for c in df if c.startswith('INDUSTRY_')]

# Add simple engagement summaries (Counts only)
eng1 = fg + ['total_engagements','unique_actions']

# Counts + Refined Recency
eng2 = eng1 + ['days_since_nonconv']

print("FG features:", fg)
print("Eng1 features:", eng1)
print("Eng3 features:", eng2)
```

# EDA & FEATURE ENGINEERING

Key features created:

- **Firmographics**: employee_count, annual_revenue, industry dummies, lead_age
- **Engagement**: total_engagements, unique_actions
- **Recency**: days_since_nonconv (excluding conversion events)

Target Variable:
- "**CONVERTED**" flag (1 if lead requested demo or attended webinar; 0 otherwise)
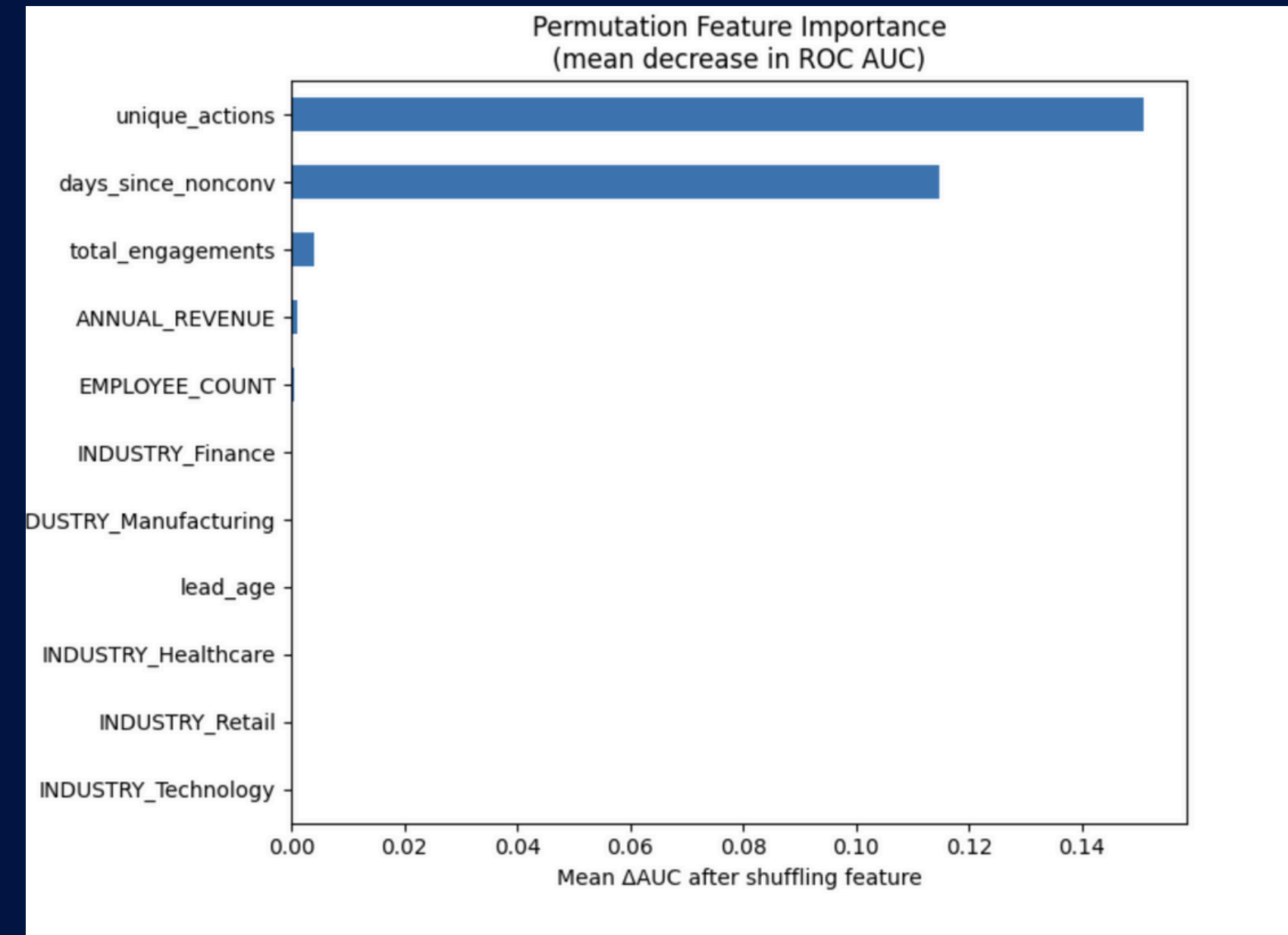
# MODEL PERFORMANCE AND EXPLAINABILITY

**Final model**: Random Forest (300 trees, balanced)

Validation:
- Time-split AUC ≈ 0.95
- 5-fold CV on eng3 features → AUC 0.99

Top drivers (Permutation Importance):
1. unique_actions
2. days_since_nonconv



Permutation Feature Importance
(mean decrease in ROC AUC)

```
Time-split ROC AUC: 0.942
              precision    recall  f1-score   support

           0       0.93      0.61      0.74        82
           1       0.88      0.98      0.93       230

    accuracy                           0.88       312
   macro avg       0.90      0.80      0.83       312
weighted avg       0.89      0.88      0.88       312
```
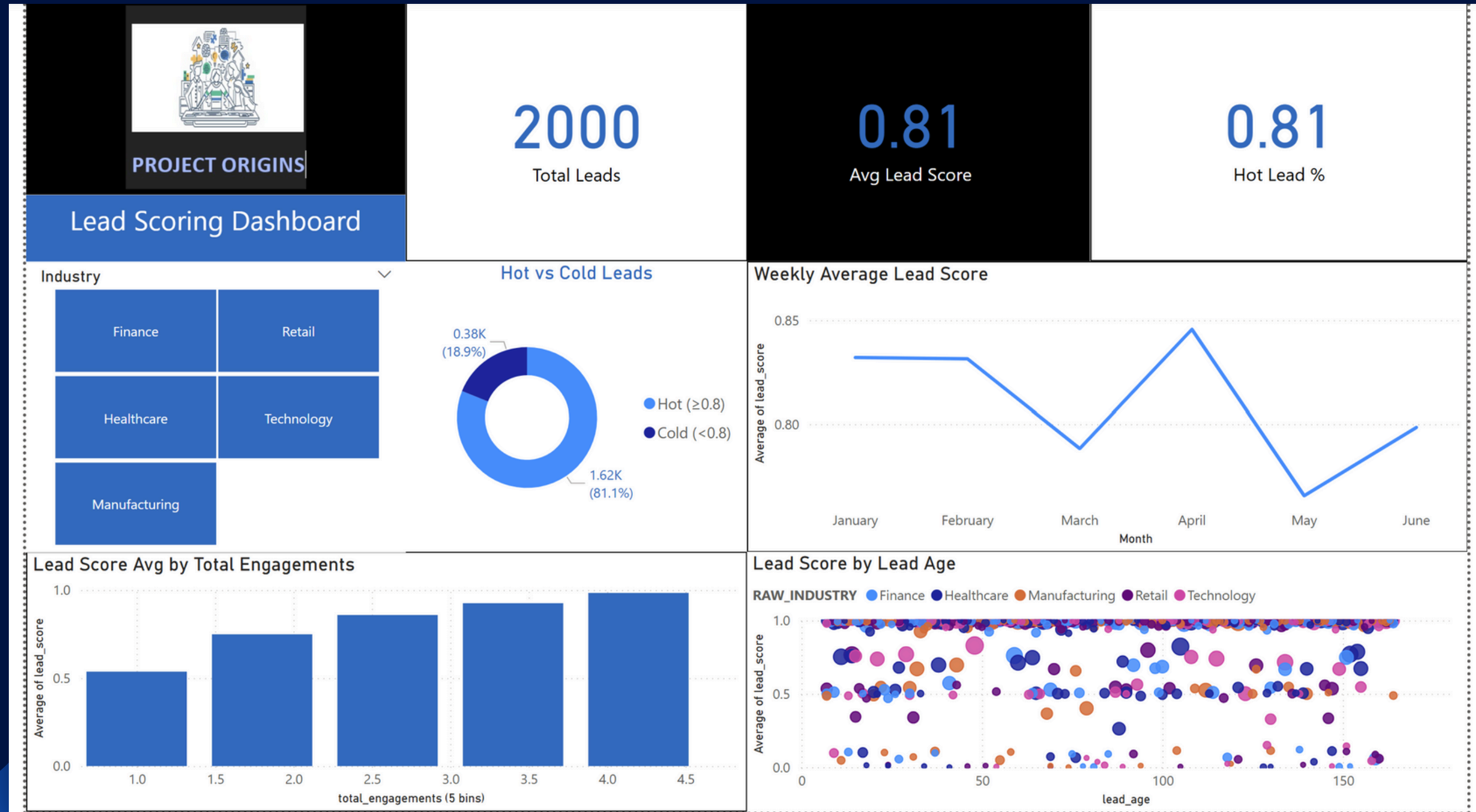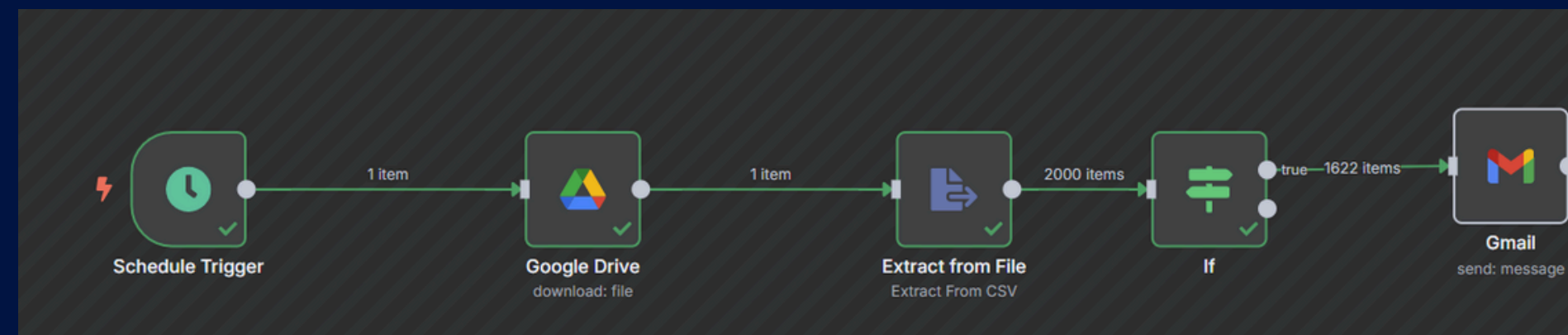
# n8n Automation & Next Steps

- n8n Workflow: Daily run → pull CSV → filter score ≥0.8 → SMTP alerts

- GTM Integration: Alerts baked into sales cadence; ICP tiers tied to outreach scripts