

# Dự báo Tỷ lệ Nghỉ việc của Nhân viên bằng Machine Learning

## Chiến lược chủ động giữ chân NS

NHÂN VIÊN PHÂN TÍCH DỮ LIỆU NHÂN SỰ

Nguyễn Đình Sang



# Nỗi đau Hiện tại - Tại sao cần quan tâm?

## Chi phí nghỉ việc

Mất kiến thức nội bộ, gián đoạn công việc

Tổng chi phí thay thế: **6-9 tháng lương**

## Tác động

Giảm tinh thần

Tăng chi phí tuyển dụng

## Tỷ lệ Turnover cao

- TB 5 năm tại KTG **15,28%**
- So với Thị trường:

*Doanh nghiệp Việt Nam: 9,6%*

*Doanh nghiệp đa QG: 6,5%*

([Báo cáo lương Năm 2024 theo Talent Net](#))

## Câu hỏi quan trọng

Liệu Kim Tín có đang bỏ lỡ cơ hội giữ chân nhân tài?

Và cách giữ chân nhân tài nếu phát hiện sớm các dấu hiệu nghỉ việc

# Giải pháp Đột phá - Dự đoán để Hành động



## Mô hình Machine Learning

Dự đoán sớm nghỉ việc nhân viên



## Mục tiêu

Can thiệp kịp thời

Thay đổi chính sách phù hợp dựa trên các yếu tố chính gây ra



# Tổng quan Dự án - Từ Dữ liệu đến Quyết định

1

Dữ liệu nhân sự (đã được chuẩn hóa)

8,702 hồ sơ toàn **KTG**, 43 cột thông tin

2

Xử lý & Lựa chọn (Giảm chiều PCA)

Input 20 cột quan trọng có ảnh hưởng đến quyết định nghỉ

3

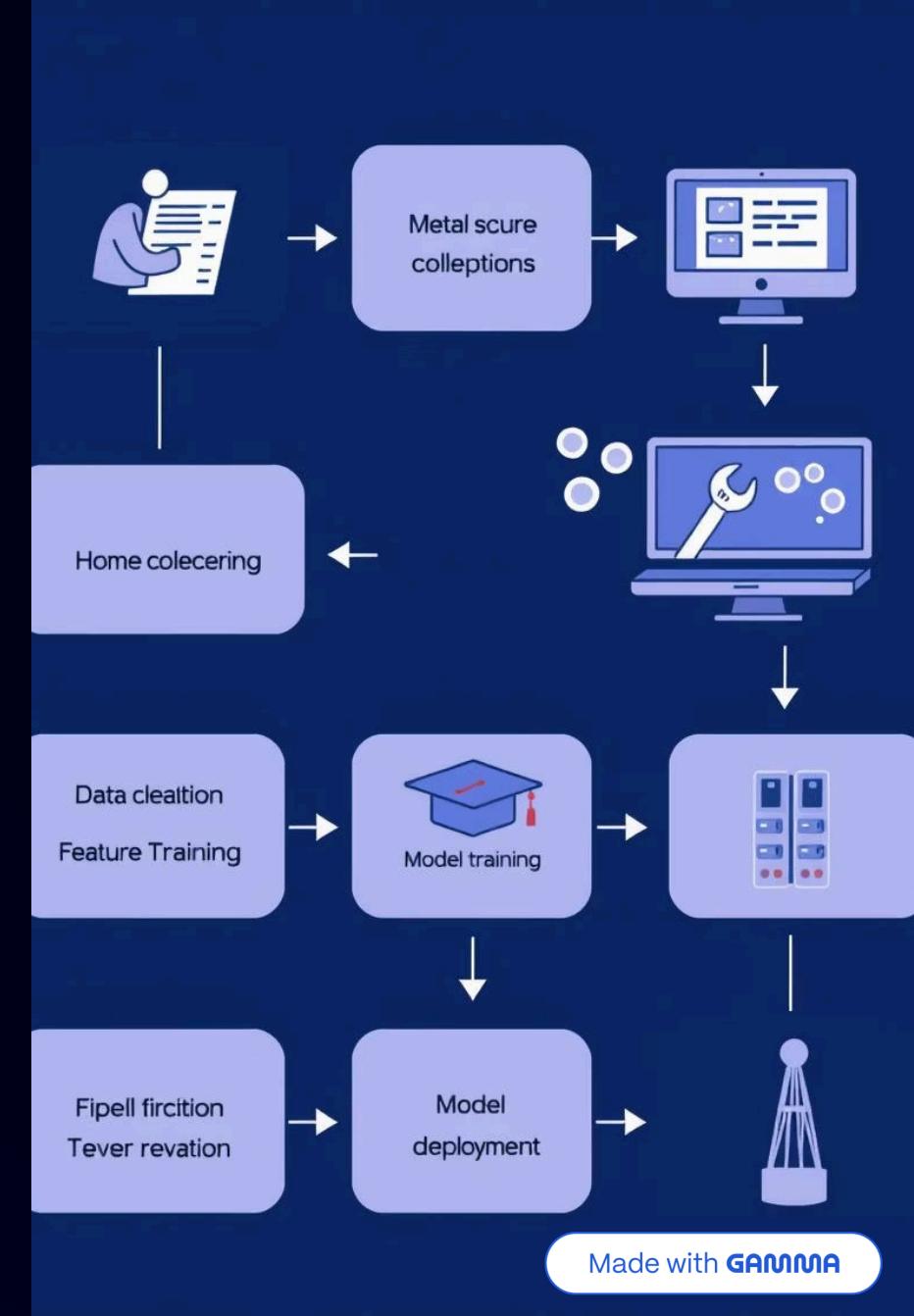
Huấn luyện lựa chọn Mô hình

Random Forest

4

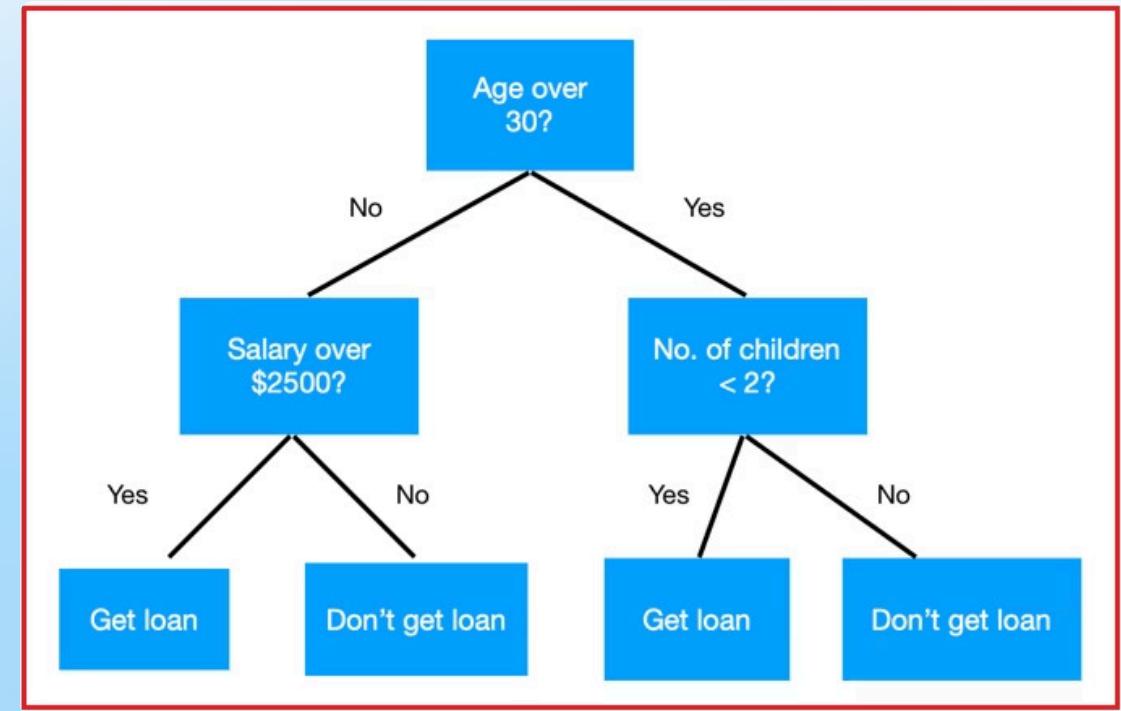
Dự đoán & Insights

Giao diện **Web** hỗ trợ HR thực hiện dự đoán



# Mô hình tối ưu nhất: Random Forest

Tổng số hàng train	6961 (80%)
<b>Tổng số hàng test</b>	<b>1741 (20%)</b>
TP (dự báo đúng nghỉ)	550
FP (dự báo sai nghỉ)	126
FN (dự báo sai ở lại)	54
TN (dự báo đúng ở lại)	1011
<b>Accuracy (% chính xác)</b>	<b>89,66%</b>





# Phát hiện Nhân viên Nguy cơ Cao (Lớp Thực tế nghỉ)

Ma trận Confuse

	Dự đoán nghỉ	Dự đoán <b>không</b> nghỉ
Thực tế nghỉ	550	54
Thực tế <b>không</b> nghỉ	126	1011

## Precision

81% trong các dự đoán là nghỉ là chính xác

$$= 550 / (550+126)$$

Precision cao đảm bảo rằng khi HR nhận được một cảnh báo, cảnh báo đó có độ tin cậy cao

## Recall

91% nhân viên nghỉ được phát hiện

$$= 550 / (550+54)$$

Recall cao để đảm bảo không bỏ sót nhiều nhân viên đang có nguy cơ nghỉ

## F1-score

86% - Chỉ số cân bằng giữa Precision và Recall.

$$F1\text{-score} = 2 * (Precision \cdot Recall) / (Precision + Recall)$$

Hiệu dụng khi là trong các bài toán có sự mất cân bằng giữa các lớp (ví dụ, số người nghỉ việc thường ít hơn số người ở lại)

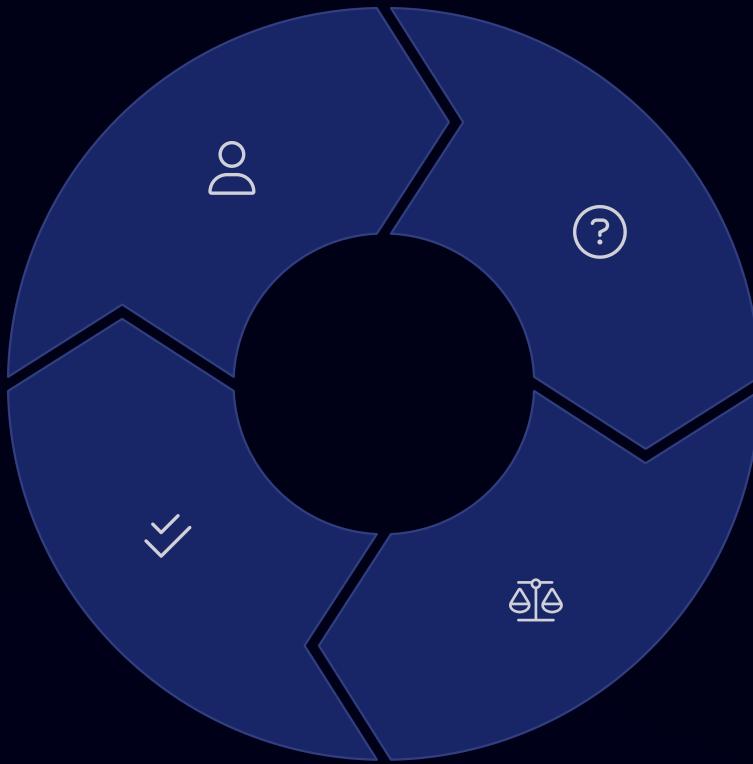
# Giải mã Thuật toán Random Forest

## Hội đồng Chuyên gia

Nhiều "cây quyết định" cùng dự đoán giúp tăng độ chính xác.

## Tổng hợp Kết quả

Dự đoán cuối cùng dựa trên đa số ý kiến của các cây quyết định.



## Chuỗi Câu hỏi Đơn giản

Mỗi cây hỏi "có/không" về thông tin nhân viên, như lương, thâm niên.

## Dự đoán Cuối cùng

Từng cây quyết định cho "Có khả năng nghỉ việc" hoặc "Ít có khả năng nghỉ".

## Sức mạnh đa dạng từ sự ngẫu nhiên

### Phân phối dữ liệu:

Mỗi chuyên gia chỉ xem ngẫu nhiên một phần dữ liệu nhân viên

=> **NHIỀU** góc nhìn

### Huấn luyện đa dạng:

Học từ các tập ngẫu nhiên các trường hợp nghỉ và ở lại

=> Độ khách quan **CAO**

# So sánh các mô hình ML sử dụng

Mô hình	Accuracy	Đánh giá
Random Forest	0.8966	Tốt nhất
Gradient Boosting	0.8817	Hiệu quả cao
Logistic Regression	0.8208	Hiệu quả kém

Random Forest nổi bật với độ chính xác cao nhất, là lựa chọn tối ưu.

# Top Đặc trưng Ảnh hưởng Nhất

Tổng Mức Lương	Thâm niên	Chức danh
Mức thưởng KPI	<i>Nguy cơ nhân viên mới và lâu năm</i>	<i>Phản ánh môi trường và cơ hội phát triển</i>
Mức thưởng Khoán	Ngày bắt đầu HĐ	Địa điểm làm việc
Lương BHXH	<i>Liên quan thời điểm onboarding trong năm hoặc thời điểm tái tục HĐLD</i>	<i>Tình thành làm việc của NS</i>
Yếu tố then chốt, cần cạnh tranh	Loại HĐ	

# Employe Attrition Predictions



Filters

Filter

	Department	Location	Location	Litter Risk	Department	Computer	High Risk
Join	Employee employees	Replomed		03/2219	6.84%	0.3%	0.3%
Prepare	Employee employees	Replomed		03/2833	6.15%	0.0%	0.3%
Supory	Employee employees	Replomed		18/2919	6.54%	0.9%	0.2%
Dirore	Pepbyer employees	Reploment		06/2229	94.8%	0.5%	0.23
Moore	Depoyer employees	Replomed		14/2219	51.8%	0.0%	0.94
Hoppe	Fambyer employees	Replomed		18/2028	61.9%	0.3%	0.37
Tenore	Leaving employest	Replomed		10/2028	24.1%	0.5%	0.3%

# Giao diện Web cho HR - Biến Dữ liệu thành Hành động

## Tiện lợi

Tải file Excel, nhận dự đoán  
nhanh

## Hỗ trợ quyết định

Danh sách nhân viên nguy cơ và  
yếu tố ảnh hưởng

## Nhanh chóng

Giúp HR can thiệp kịp thời, hiệu quả

Demo: [LINK](#)



# Lợi ích Thực tiễn cho Doanh nghiệp

1

## Tiết kiệm chi phí

Giảm chi phí tuyển dụng nhờ giữ chân nhân viên quan trọng.

2

## Nâng cao hiệu suất

Giữ đội ngũ kinh nghiệm, giảm gián đoạn công việc.

3

## Cải thiện văn hóa

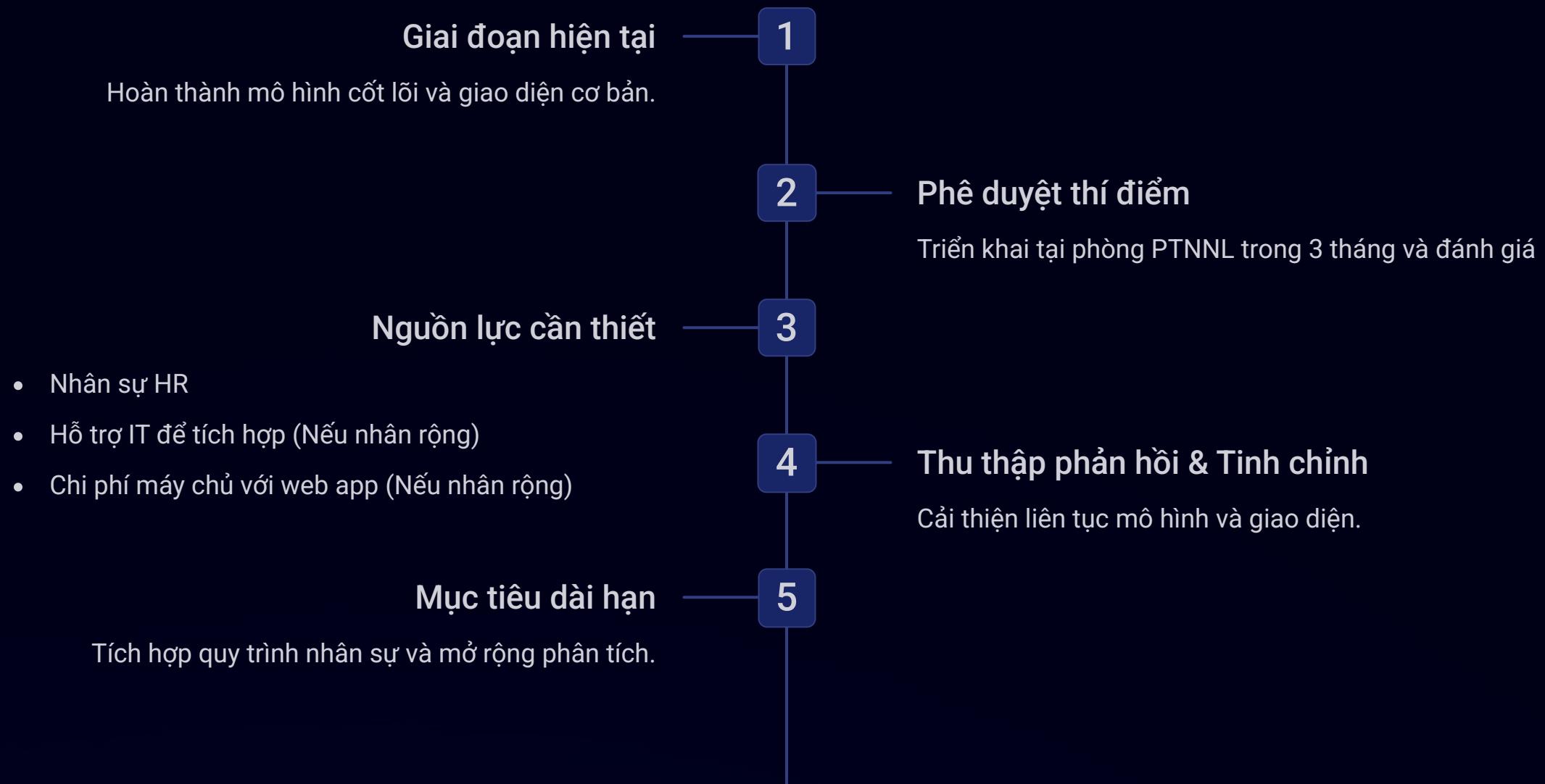
Thể hiện sự quan tâm chủ động tới nhân viên.

4

## Quyết định thông minh

Dựa trên dữ liệu, hạn chế cảm tính chủ quan.

# Lộ trình & Đề xuất Triển khai



# THANK YOU