

ĐẠI HỌC UEH - TRƯỜNG CÔNG NGHỆ VÀ THIẾT KẾ
KHOA CÔNG NGHỆ THÔNG TIN KINH DOANH



KHÓA LUẬN TỐT NGHIỆP

GVHD: TS. Huỳnh Văn Đức

Tên đề tài: Website ứng dụng khoa học dữ liệu trong lĩnh vực quản lý nhân sự

Sinh viên: Nguyễn Thị Thảo Ly

MSSV: 31201024417

K46-ST001

Tp. Hồ Chí Minh, Tháng 10/ 2023

Mục lục

Mục lục.....	2
TÓM TẮT.....	4
I. TỔNG QUAN.....	5
1. Yêu cầu.....	5
2. Phạm vi (Scope of the Project).....	5
3. Công nghệ sử dụng (Technology Stack).....	5
4. Dashboard trong lĩnh vực quản trị nhân sự và các thông tin được sử dụng.....	6
5. Về hai model được sử dụng.....	7
5.1. Model Predicting_Promotion.....	7
5.2. Model Predicting_Retention.....	8
II. Xây dựng hệ thống.....	10
1. Khái quát hệ thống.....	10
1.1.1 Kiến trúc tổng thể Data Warehouse.....	10
1.1.2. Biểu đồ thành phần.....	11
2. Data warehousing với Snowflake và Data Build Tool.....	12
1.1. Tạo kho dữ liệu trên Snowflake.....	12
1.1.1. DIM_DEPARTMENT.....	12
1.1.2. DIM_EDULEVEL.....	12
1.1.3 DIM_EMPLOYEE.....	13
1.1.3. DIM_QUESTION_SURVEY_EMPLOYEES và ENGAGEMENT.....	14
1.1.4. DIM_QUESTION_SURVEY_MANAGER và EMPLOYEE_PERFORMANCE_SURVEY.....	15
1.1.5. DIM_WORKSCHE.....	16
1.1.6. EMPLOYEEKPI.....	16
1.1.7. ATTRITION.....	17
1.1.8. DEFECT.....	17
1.1.9. ABSENTEEISM.....	19
1.1.10. SKILL_RECORD.....	20
1.2. Xây dựng Model bằng Data Build tool.....	21
Linear Graph workforce.....	24
Toàn bộ mô hình dữ liệu tuyến tính (linear-graph).....	24
Linear Graph cho chỉ số OLE -Dashboard Performance.....	25
RETENTION cung cấp các feature cần thiết để dự đoán khả năng nghỉ việc.....	26

1.3. Data Model.....	27
2. Phân tích hệ thống.....	28
2.1. Use Case Diagram.....	28
2.2. Activity Diagram.....	29
2.2.1 Xem Dashboard.....	30
2.2.2. Dự đoán khả năng thăng tiến.....	31
2.2.3. Dự đoán khả năng nghỉ việc.....	33
2.3 Sequence Diagram.....	34
2.2.1. Dự đoán khả năng thăng tiến.....	34
2.2.2. Dự đoán khả năng nghỉ việc.....	35
3 .Giao diện và Dashboard.....	36
3.1. Dashboard Workforce Management.....	36
3.2. Dashboard Performance.....	37
3.4. Dashboard Talent Management.....	38
3.5. Dự đoán khả năng thăng chức với dữ liệu được nhập từ người dùng.....	39
3.6. Thêm giá trị mới cho tranining dataset.....	40
3.7. Dự đoán trên danh sách nhân viên được chọn.....	40
3.8 Dự đoán và đề xuất nhân viên trên mỗi phòng ban.....	42
3.9. Dự đoán nhân viên có xác suất nghỉ việc cao.....	43
III. Tổng kết.....	45
1. Kiến thức tìm hiểu.....	45
1.1. Streamlit.....	45
1.2. Snowflake và Data Build Tool.....	45
1.3. Về lĩnh vực quản trị nhân sự.....	45
1.4. Kết hợp website với các yếu tố khoa học dữ liệu.....	46
2. Hướng phát triển.....	47
3. Tài liệu tham khảo.....	47

TÓM TẮT

Thấy được tầm quan trọng của bộ phận quản trị nhân sự đối với doanh nghiệp, dự án này tập trung xây dựng website phân tích dữ liệu, ứng dụng các yếu tố của Data Science trong lĩnh vực quản trị nhân sự để giúp các nhà quản lý kịp thời ra các quyết định quan trọng, khai thác tiềm năng về nguồn nhân lực. Từ đó giúp giảm các chi phí ẩn do việc thay đổi nhân sự, tạo ra môi trường có thể cộng tác lâu dài với nhân viên.

Dự án sẽ sử dụng 2 mô hình machine learning để dự đoán khả năng thăng tiến, và khả năng nghỉ việc của nhân viên, đồng thời diễn tả các trường thông tin quan trọng trong lĩnh vực quản trị nhân sự nói chung và các yếu tố liên quan tới kết quả dự đoán nói riêng, một yếu tố quan trọng là khả năng cập nhập dữ liệu theo thời gian thực. Dự án sẽ đi từ bước thu thập data từ S3 bucket để tạo kho chứa raw data trên Snowflake, sau đó xây dựng logic dữ liệu tạo Data Warehouse và Data Mart, cuối cùng biểu đồ hóa xây dựng dashboard và chức năng dự đoán trên nền tảng Website.

Phần cuối bài báo các đề cập tới kiến thức và quá trình lựa chọn, tìm hiểu các lĩnh vực và công cụ cần thiết để hoàn thành dự án, cũng như tổng kết về các hạn chế cũng như hướng phát triển của dự án trong tương lai.

I. TỔNG QUAN

1. Yêu cầu

Hệ thống của dự án đáp ứng các yêu cầu chức năng sau đây:

- Tạo Dashboard cho một số thông tin quan trọng trong lĩnh vực quản trị nhân sự
- Dự đoán nhân viên nào có tiềm năng để cân nhắc khen thưởng hoặc thăng chức
 - Dự đoán trên các đối tượng cụ thể
 - Có thể thêm mới dữ liệu, retrain model tăng độ hiệu quả dự đoán trong tương lai
- Dự đoán khả năng nghỉ việc của nhân viên

2. Phạm vi (Scope of the Project)

Dự án này bao gồm việc thiết kế và phát triển ứng dụng, từ việc xây dựng giao diện người dùng cho đến việc triển khai các mô hình machine learning. Đối với Dashboard, sẽ có sự tham khảo tìm hiểu về kiến thức quản trị nhân sự cũng như các trường thông tin quan trọng thường được trình bày trên Dashboard. Đối với dữ liệu, các bảng và cột được quyết định dựa trên thông tin cần có của Dashboard, dữ liệu của nhân viên dùng để thêm vào các bảng được thu thập một phần từ các nguồn mở, một phần khác tạo giả dữ liệu dựa trên các dataset có sẵn cho mô hình dự đoán, để từ đó giảm thiểu tính sai lệch của kết quả dự đoán. Ngoài ra, em cũng tập trung vào việc đảm bảo rằng ứng dụng hoạt động một cách mượt mà và trực quan, bảo đảm trải nghiệm người dùng.

3. Công nghệ sử dụng (Technology Stack)

Dự án của em sử dụng các công nghệ và thư viện sau để đảm bảo hiệu suất và tính năng tốt nhất:

- Streamlit: frontend
- FastAPI: backend
- Snowflake: Snowflake được sử dụng như một kho dữ liệu đám mây để lưu trữ và quản lý dữ liệu.
- DBT (data build tool): dbt được sử dụng để biến đổi và chuẩn bị dữ liệu trong Snowflake. Nó giúp xây dựng, kiểm tra và triển khai mã SQL trong kho dữ liệu.
- Mô hình Machine Learning từ bên thứ ba: Em đã chuẩn bị hai mô hình machine learning có sẵn từ bên thứ 3 để sử dụng trong ứng dụng. Một mô hình được sử dụng để dự đoán tỷ lệ nghỉ việc của nhân viên và mô hình kia được sử dụng để dự đoán khả năng thăng tiến.
- Ngôn ngữ lập trình chính: Python
- Khác: s3 bucket của aws để chứa dữ liệu tạm thời được tải từ hệ thống local

4. Dashboard trong lĩnh vực quản trị nhân sự và các thông tin được sử dụng

Các công ty thành công trong việc tận dụng phân tích dữ liệu nhân sự sẽ có vị trí vượt trội so với các đối thủ cạnh tranh của họ khi nói đến chiến lược quản lý nhân tài. Do đó, việc biết ai đang làm việc trong công ty của bạn và họ làm việc như thế nào là rất quan trọng.

Dashboard trong bài project này sẽ thể hiện một số thông tin như sau

1. **Số liệu tổng quát của lực lượng lao động:** Sự thay đổi của nhân viên toàn thời gian so với bán thời gian là một chỉ số về tính hiệu quả (KPI), lấy ví dụ khi số lượng nhân viên toàn thời gian giảm đáng kể trong khi số lượng nhân viên bán thời gian tăng mạnh. Có nhiều lý do có thể giải thích sự thay đổi này, lý do chính là khó khăn tài chính, hoặc giảm lượng công việc cần ít người làm việc 40 giờ một tuần. Nhìn vào hành vi của nhân viên trong năm đầu tiên làm việc, điều nổi bật là xu hướng rời bỏ vị trí sau khoảng một năm đến 3 năm đầu tiên. Một số thông tin khác bao gồm tỷ lệ nam nữ trong công việc, để các nhà quản lý bảo đảm tạo ra môi trường làm việc giúp nhân viên thích ứng tốt
2. **Hành vi và hiệu suất của nhân viên:** Nó cung cấp cho các quản lý cái nhìn tổng quan về lực lượng lao động và cách họ làm việc: phần đầu tiên cho chúng ta biết về tình trạng nghỉ phép, số ngày nghỉ phép trung bình trong một năm, tỷ lệ trung bình và sự thay đổi của chúng trong khoảng thời gian qua. Đây là một chỉ số quan trọng minh họa cho sự tận tâm và động lực của người lao động. Phần thứ hai của này xem xét Hiệu quả Lao động Tổng thể (OLE), tức là năng suất lao động. Nó giúp đo lường hiệu quả lao động theo thời gian theo các tiêu chí và mục tiêu cụ thể mà doanh nghiệp đặt ra.
3. **Quản lý tài năng:** đưa ra các số liệu về tổng quan nhân lực và chi phí lương hàng tháng cho toàn bộ công ty, tiếp theo đưa ra cái nhìn về tỉ lệ thay đổi nhân sự giữa tình nguyện và bị sa thải trên phòng ban và trên số năm làm việc tại công ty. Tại phần này, còn trình bày về phần trăm nhân sự cảm thấy hòa nhập với môi trường làm việc và vị trí trí công việc theo số năm đồng hành với doanh nghiệp. Phần cuối cùng thể hiện điểm đánh giá của nhân viên theo số năm gắn bó với công ty

5. Về hai model được sử dụng

5.1. Model Predicting_Promotion

Chi tiết về dataset dùng để train model

RangeIndex: 54808 entries, 0 to 54807

Data columns (total 14 columns):

#	Column	Non-Null Count	Dtype
0	employee_id	54808 non-null	int64
1	department	54808 non-null	object
2	region	54808 non-null	object
3	education	52399 non-null	object
4	gender	54808 non-null	object
5	recruitment_channel	54808 non-null	object
6	no_of_trainings	54808 non-null	int64
7	age	54808 non-null	int64
8	previous_year_rating	50684 non-null	float64
9	length_of_service	54808 non-null	int64
10	KPIs_met >80%	54808 non-null	int64
11	awards_won?	54808 non-null	int64
12	avg_training_score	54808 non-null	int64
13	is_promoted	54808 non-null	int64

Dựa trên kết quả của mô hình dự đoán, các biến sau đây có mối tương quan với việc được thăng chức:

Phòng ban (Department): Có mối tương quan với việc được thăng chức (P-Value: 1.96). Điều này có thể cho thấy rằng một số phòng ban có tỷ lệ thăng chức cao hơn so với những phòng ban khác.

Khu vực (Region): Có mối tương quan với việc được thăng chức (P-Value: 0.0). Điều này có thể cho thấy rằng một số khu vực có tỷ lệ thăng chức cao hơn so với những khu vực khác.

Trình độ học vấn (Education): Có mối tương quan với việc được thăng chức (P-Value: 1.45). Điều này có thể cho thấy rằng những người có trình độ học vấn cao hơn có khả năng được thăng chức cao hơn.

Giới tính (Gender): Có mối tương quan với việc được thăng chức (P-Value: 1.94). Điều này có thể cho thấy rằng giới tính có ảnh hưởng đến khả năng được thăng chức.

Kênh tuyển dụng (Recruitment Channel): Có mối tương quan với việc được thăng chức (P-Value: 1.85). Điều này có thể cho thấy rằng một số kênh tuyển dụng mang lại tỷ lệ thăng chức cao hơn so với những kênh khác.

Số lượng đào tạo (No of Trainings): Có mối tương quan với việc được thăng chức (P-Value: 6.38). Điều này có thể cho thấy rằng những người nhận được nhiều đào tạo hơn có khả năng được thăng chức cao hơn.

Đánh giá năm trước (Previous Year Rating): Có mối tương quan với việc được thăng chức (P-Value: 0.0). Điều này cho thấy rằng những người nhận được đánh giá cao trong năm trước có khả năng được thăng chức cao hơn.

Thời gian làm việc (Length of Service): Có mối tương quan với việc được thăng chức (P-Value: 2.933826971815107e-59). Điều này cho thấy rằng những người làm việc lâu hơn có khả năng được thăng chức cao hơn.

Đạt KPIs >80%: Có mối tương quan với việc được thăng chức (P-Value: 0.0). Điều này cho biết rằng những người đạt KPIs >80% có khả năng được thăng chức cao hơn.

Model sử dụng Xgboost để tiến hành dự đoán, độ chính xác của model

	precision	recall	f1-score	support
0	0.98	0.86	0.91	15127
1	0.87	0.98	0.92	14957
accuracy			0.92	30084
macro avg	0.92	0.92	0.92	30084
weighted avg	0.92	0.92	0.92	30084

5.2. Model Predicting Retention

Sau khi chạy các bước về phân tích dữ liệu, biểu đồ chỉ ra được:

- Tuổi: Trung bình tuổi của những người ở lại cao hơn so với những người đã rời đi.
- Mức lương hàng ngày: Mức lương hàng ngày của những người ở lại cao hơn.
- Khoảng cách từ nhà đến nơi làm việc: Những người ở lại thường sống gần nơi làm việc hơn.
- Mức độ hài lòng với môi trường làm việc và Mức độ hài lòng với công việc**: Những người ở lại thường hài lòng hơn với công việc và môi trường làm việc.
- Mức độ lựa chọn cổ phiếu: Những người ở lại thường có mức lựa chọn cổ phiếu cao hơn.

Chi tiết về dataset dùng để train model

RangeIndex: 1470 entries, 0 to 1469

Data columns (total 35 columns):

#	Column	Non-Null Count	Dtype
0	Age	1470 non-null	int64
1	Attrition	1470 non-null	object
2	BusinessTravel	1470 non-null	object
3	DailyRate	1470 non-null	int64
4	Department	1470 non-null	object
5	DistanceFromHome	1470 non-null	int64
6	Education	1470 non-null	int64
7	EducationField	1470 non-null	object
8	EmployeeCount	1470 non-null	int64
9	EmployeeNumber	1470 non-null	int64
10	EnvironmentSatisfaction	1470 non-null	int64
11	Gender	1470 non-null	object
12	HourlyRate	1470 non-null	int64
13	JobInvolvement	1470 non-null	int64
14	JobLevel	1470 non-null	int64
15	JobRole	1470 non-null	object
16	JobSatisfaction	1470 non-null	int64
17	MaritalStatus	1470 non-null	object
18	MonthlyIncome	1470 non-null	int64
19	MonthlyRate	1470 non-null	int64
20	NumCompaniesWorked	1470 non-null	int64
21	Over18	1470 non-null	object
22	OverTime	1470 non-null	object
23	PercentSalaryHike	1470 non-null	int64
24	PerformanceRating	1470 non-null	int64
25	RelationshipSatisfaction	1470 non-null	int64
26	StandardHours	1470 non-null	int64
27	StockOptionLevel	1470 non-null	int64
28	TotalWorkingYears	1470 non-null	int64
29	TrainingTimesLastYear	1470 non-null	int64
30	WorkLifeBalance	1470 non-null	int64
31	YearsAtCompany	1470 non-null	int64
32	YearsInCurrentRole	1470 non-null	int64
33	YearsSinceLastPromotion	1470 non-null	int64
34	YearsWithCurrManager	1470 non-null	int64

Những nhận xét này cho thấy rằng các yếu tố như tuổi, mức lương hàng ngày, khoảng cách từ nhà

đến nơi làm việc, mức độ hài lòng với công việc và môi trường làm việc, cũng như mức độ lựa chọn cổ phiếu có thể ảnh hưởng đến quyết định của một nhân viên có nên ở lại hay rời bỏ công ty. Model sử dụng Logistic Regression để tiến hành dự đoán, kết quả sau khi chạy:

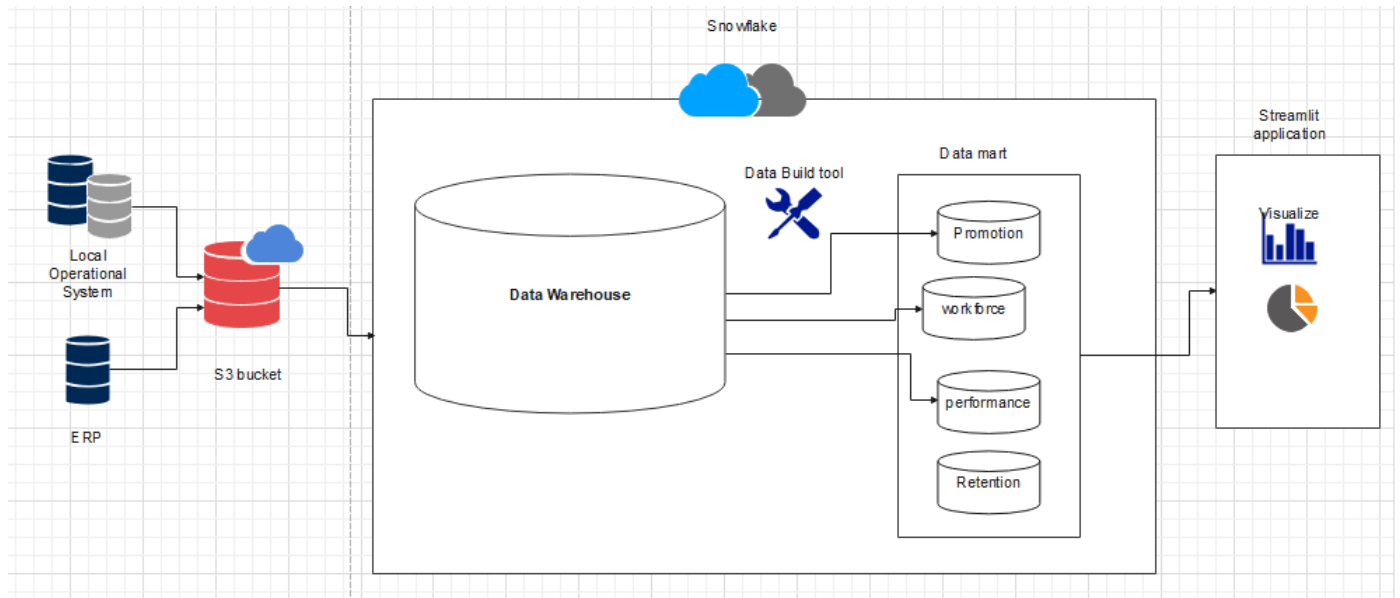
	precision	recall	f1-score	support
0	0.87	0.96	0.92	310
1	0.56	0.26	0.35	58
accuracy			0.85	368
macro avg	0.71	0.61	0.63	368
weighted avg	0.82	0.85	0.83	368

II. Xây dựng hệ thống

1. Khái quát hệ thống

1.1.1 Kiến trúc tổng thể Data Warehouse

Kho dữ liệu (DW hoặc DWH) là một hệ thống phức tạp lưu trữ dữ liệu được sử dụng cho việc dự báo, báo cáo và phân tích dữ liệu. Nó liên quan đến việc thu thập, làm sạch và biến đổi dữ liệu từ các luồng dữ liệu khác nhau và tải nó vào các bảng Dimensiona/Fact

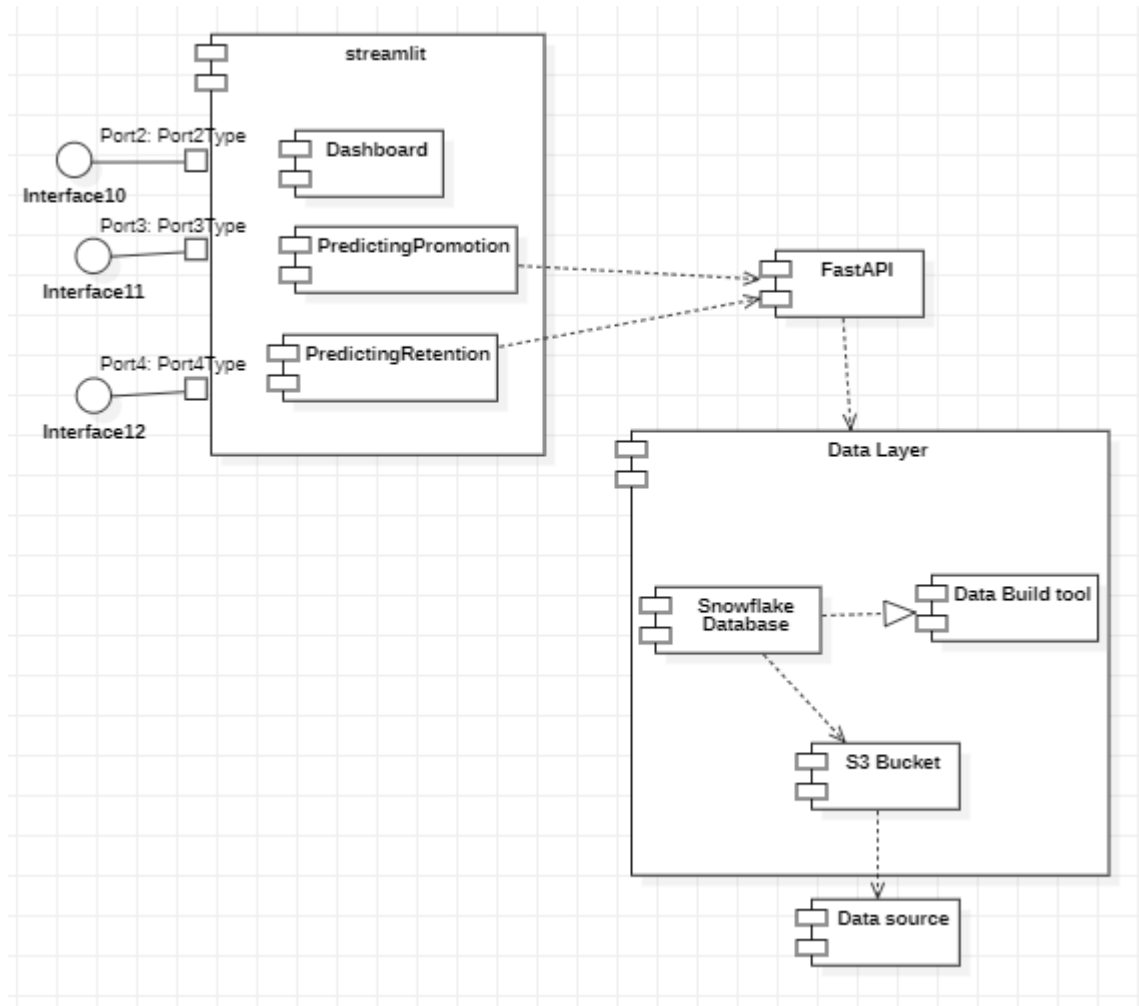


Kiến trúc tổng thể

1.1.2. Biểu đồ thành phần

Trong hệ thống sẽ có các thành phần sau tương tác với nhau:

- Streamlit: được sử dụng như một lớp frontend, nơi người dùng có thể tương tác với ứng dụng.
- FastAPI: được sử dụng như một lớp backend, nơi nó xử lý các yêu cầu từ frontend (Streamlit), truy vấn nguồn dữ liệu và trả về kết quả.
- Nguồn Dữ liệu (Data Layer): đây là nơi lưu trữ và quản lý dữ liệu. FastAPI sẽ truy vấn nguồn dữ liệu này để lấy thông tin cần thiết để xử lý các yêu cầu từ frontend.







2. Data warehousing với Snowflake và Data Build Tool

1.1. Tạo kho dữ liệu trên Snowflake

1.1.1. DIM_DEPARTMENT

Bảng dữ liệu giải thích nghĩa của mã phòng ban với:





- CODE: mã phòng ban
- DEPARTMENT: tên phòng ban

NAME ↑	TYPE
 CODE	 Varchar
 DEPARTMENT	 Varchar

1.1.2. DIM_EDULEVEL

Bảng dữ liệu giải thích nghĩa trình độ học vấn với:

- LEVELNUMBER: đánh dấu cấp độ theo 1, 2, 3, 4, 5
- EXPLANATION: bằng cấp tương đương theo từng cấp độ

NAME ↓	TYPE
 LEVELNUMBER	 Number
 EXPLANATION	 Varchar

1.1.3 DIM_EMPLOYEE

Bảng tổng hợp thông tin căn bản của nhân viên, sử dụng tạo dashboard Workforce, và Performance và trích xuất dữ liệu của nhân viên sau khi dự đoán khả năng thăng tiến và khả năng nghỉ việc.







NAME ↑	TYPE
<div><div></div>AGE</div>	# Number
<div><div></div>DEPARTMENT_ID</div>	<u>A</u> Varchar
<div><div></div>EDUCATION</div>	<u>A</u> Varchar
<div><div></div>EMPLOYEE_NAME</div>	<u>A</u> Varchar
<div><div></div>EMPLOYEE_NUMBER</div>	<u>A</u> Varchar
<div><div></div>GENDER</div>	<u>A</u> Varchar
<div><div></div>HIREDATE</div>	🕒 Timestamp_NTZ
<div><div></div>JOB_ROLE</div>	<u>A</u> Varchar
<div><div></div>MARITAL_STATUS</div>	<u>A</u> Varchar
<div><div></div>NUM</div>	# Number
<div><div></div>PAYROLLID</div>	<u>A</u> Varchar
<div><div></div>RECRUITMENT_CHANNEL</div>	<u>A</u> Varchar
<div><div></div>REGION</div>	<u>A</u> Varchar
<div><div></div>STATUS</div>	# Number
<div><div></div>WORKSCHEDULE</div>	<u>A</u> Varchar

1.1.3. DIM_QUESTION_SURVEY_EMPLOYEES và ENGAGEMENT

2 bảng tổng hợp dữ liệu liên quan tới khả năng thích ứng của nhân viên trong môi trường làm việc và đối với vị trí công việc họ đảm nhận. Cần cho chức năng dự đoán nghỉ việc, các thông tin liên quan tới độ lành mạnh của môi trường làm việc được thể hiện trên Dashboard.

- DIM_QUESTION: FACTOR gồm các khía cạnh như (mối quan hệ -relationship, cân bằng công việc và cuộc sống-Worklife Balance, thỏa mãn về vị trí công việc -job satisfaction,...NAME gồm đối tượng các câu hỏi khảo sát hướng tới, QUESTION các câu hỏi trong form khảo sát.
- ENGAGEMENT: ghi lại dữ liệu điền form của nhân viên, ngày tháng thực hiện khảo sát và điểm số đánh giá cho từng câu hỏi

Bảng DIM_QUESTION

NAME ↑	TYPE
 FACTOR	<u>A</u> Varchar
 FACTORID	<u>A</u> Varchar
 NAME	<u>A</u> Varchar
 QUESTION	<u>A</u> Varchar
 QUESTIONID	<u>A</u> Varchar
 SCOREID	<u>A</u> Varchar

Bảng *ENGAGEMENT*

NAME ↑	TYPE
DATE_	Timestamp_NTZ
EMPLOYEE_ID	Number
FACTORID	Varchar
QUESTIONID	Varchar
SCOREID	Varchar
VOTING	Number

1.1.4. DIM_QUESTION_SURVEY_MANAGER và EMPLOYEE_PERFORMANCE_SURVEY

2 Bảng lưu dữ liệu đánh giá của quản lý/ trường phòng ban đối với từng nhân viên theo bộ các câu hỏi cho từng nhóm kỹ năng khác nhau. Bảng cung cấp thông tin cần cho dự đoán thăng tiến

- DIM_QUESTION_PERFORMANCE: giải thích ý nghĩa của từng mã câu hỏi, và cho kỹ năng nào
- STG_PERFORMANCE_EVL : lưu dữ liệu đánh giá







NAME ↓	TYPE
TYPE	Varchar
QUESTIONCODE	Varchar
EXPLANATION	Varchar

NAME ↑	TYPE
EMPLOYEE_ID	Number
PERFORMANCERATING	Number
QUESTIONCODE	Varchar

1.1.5. DIM_WORKSCHE

Bảng mô tả ý nghĩa của từng mã kiểu công việc (Full -time, part time, Seasonal, ...), số giờ làm việc tối thiểu, số ngày làm việc trong tháng

- WORKSCH: Work Schedule
- WORKSCHT: Work Schedule Translation
- WSTYP: Work Schedule Type
- WSTYPT: Work Schedule Type Translation
- HOURSINDAY: số giờ làm việc trong một ngày
- WORKDAYSINMONTH: số ngày làm việc trong một tháng

NAME ↑	TYPE
 HOURSINDAY	# Number
 WORKDAYSINMONTH	# Number
 WORKSCH	<u>A</u> Varchar
 WORKSCHT	<u>A</u> Varchar
 WSTYP	<u>A</u> Varchar
 WSTYPT	<u>A</u> Varchar

1.1.6. EMPLOYEEKPI

Bảng ghi lại dữ liệu về KPI của từng nhân viên, nhằm cung cấp thông tin cần cho dự đoán tăng tiến và các trường dữ liệu cho Dashboard, bảng cho biết:

- EMPLOYEENUMBER: mã nhân viên
- NUM: mã KPI
- GOAL và ACTUAL_OUTPUT: mục tiêu và kết quả thực tế
- KPI: phần trăm hoàn thành

NAME ↑	TYPE
ACTUAL_OUTPUT	# Number
CODE	A Varchar
EMPLOYEENUMBER	# Number
GOAL	# Number
KPI	# Float
MONTH_	# Number
NUM	# Number

1.1.7. ATTRITION

Ghi lại dữ liệu liên quan đến nhân viên nghỉ việc, cung cấp các thông tin về mã nhân viên, ngày nghỉ việc và loại nghỉ việc- EMPLOYEESTATUS (tình nguyện, bị buộc thôi việc,..)

NAME ↑	TYPE
ATTRITIONID	# Number
DATEOFTERMINATION	🕒 Date
EMPID	# Number
EMPLOYMENTSTATUS	A Varchar

1.1.8. DEFECT

Ghi lại dữ liệu về các sản phẩm bị lỗi hay không đạt chuẩn, dữ liệu này là một phần để tính chỉ số OLE (overall labor effectiveness)


NAME ↑

TYPE

CODE	Varchar
EMPLOYEENUMBER	Number
MONTH_	Number
NUMBER_OF_DEFECTS	Number
YEAR_	Number




























1.1.9. ABSENTEEISM

Bảng ghi lại những lần vắng mặt của nhân viên, bao gồm có lý do chính đáng và không chính đáng. Bảng cho biết ngày và giờ ghi lại vắng mặt và lý do vắng mặt. Bảng cần thiết cho việc tính toán chỉ số OLE cùng với EMPLOYEEKPI và DEFECT

NAME 	TYPE
 ABSENCE_REASON	 Varchar
 DATE	 Date
 EMPLOYEE_ID	 Number
 EMPLOYEE_NAME	 Varchar
 NUM	 Number
 TIME_IN	 Time
 TIME_OUT	 Time

1.1.10. SKILL_RECORD

Bảng ghi lại danh sách kỹ năng của nhân viên

NAME 	TYPE
 ALE	 Varchar
 AWARDS	 Varchar
 BAPI	 Varchar
 BRF	 Varchar
 DIALOG_PROGRAMMING	 Varchar
 EMPLOYEENUMBER	 Number
 FLOORPLAN_MANAGER	 Varchar
 SAPSCRIPT	 Varchar
 SAP_ABAP	 Varchar
 SAP_ALV	 Varchar
 SAP_BADI	 Varchar
 SAP_PI	 Varchar
 USER_EXIT	 Varchar

1.2. Xây dựng Model bằng Data Build tool

Sau khi có các bảng raw data trên Snowflake, việc tiếp theo là kết nối snowflake với dbt cloud để xây dựng logic dữ liệu. Ví dụ, chúng ta có code của một trong số bảng chính là WORKFORCE -dùng để lưu trữ dữ liệu liên quan tới cơ cấu nhân sự và các thông tin căn bản của nhân viên sau khi giải nghĩa các mã như mã phòng ban, mã payroll,...

```
1  {{ config(materialized="table") }}
2
3  with
4      employees as (
5          select
6              *,
7              case
8                  when gender = 1 then 'Female' when gender = 0 then 'Male' else 'Unknown'
9              end as gender_
10             from {{ ref("dim_ems") }}
11
12     ),
13     department as (select * from {{ ref("dim_department") }}),
14     salary_range as (select * from {{ ref("dim_sal") }}),
15     payroll_rec as (select * from {{ ref("payroll") }}),
16     base as (
17         select
18             *,
19             cast(
20                 replace(
21                     replace(split_part(salary_range.sallvlt, ' - ', 1), '$', ''),
22                     ',', ''
23                 ) as int
24             ) as fromsal,
25             cast(
26                 replace(
27                     replace(split_part(salary_range.sallvlt, ' - ', 2), '$', ''),
28                     ',', ''
29                 ) as int
30             ) as tosal
31         from salary_range
32     ),
33     payroll_range as (
```

```

35 payroll_range as (
36     select payroll_rec.*, base.salary_level as salary_level, base.sallvlt
37     from payroll_rec
38     join base on payroll_rec.annualincome between base.fromsal and base.tosal
39 ),
40 education as (select * from {{ ref("dim_edulevel") }}),
41 workschedule as (select * from {{ ref("dim_worksche") }}),
42 workforce_manage as (
43     select
44         employees.employee_id as employee_id,
45         employees.name,
46         employees.age as Age,
47         department.department_explain as Department,
48         employees.maritalstatus as maritalstatus,
49         employees.gender_ as gender,
50         employees.hiredate,
51         payroll_range.annualincome as Annual_Income,
52         payroll_range.sallvlt as salary_range,
53         education.explanation as education,
54         workschedule.workschedule_translation,
55         employees.education_level as education_level,
56         employees.Jobrole as Jobrole,
57         employees.status as status
58     from employees
59     left join payroll_range on (employees.payrollid = payroll_range.payrollid)
60     left join department on (employees.departmentid = department.departmentid)
61     left join education using (education_level)
62     left join workschedule using (workschedule)
63 )
64 select *
65 from workforce_manage
66

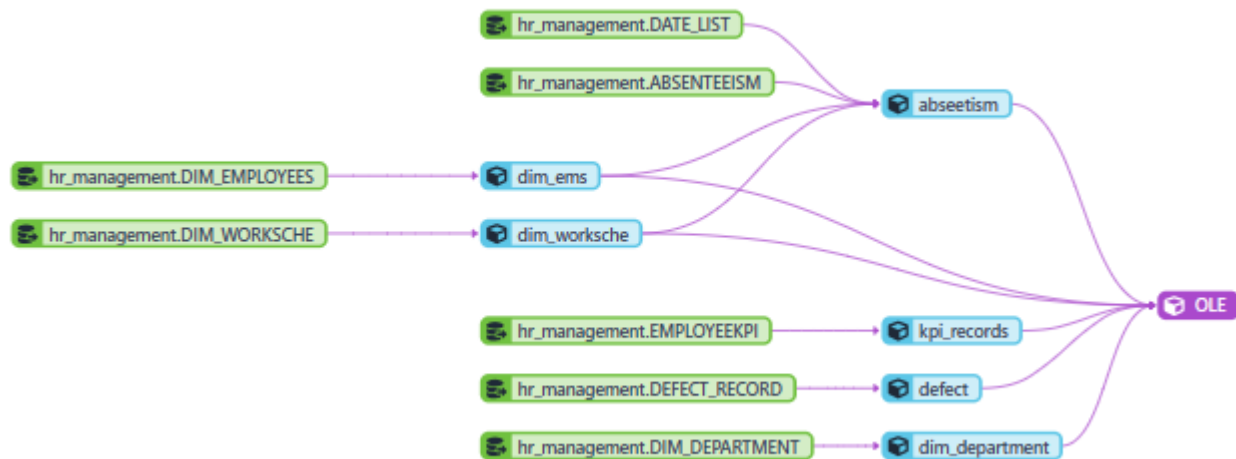
```

Kết quả ta có WORKFORCE, và biểu đồ tuyến tính thể hiện đường đi logic của dữ liệu như sau

workforce table	
Details	Description
Columns	Referenced By
Depends On	Code
Columns	
COLUMN	TYPE
employee_id	TEXT
name	TEXT
age	NUMBER
department	TEXT
maritalstatus	TEXT
gender	TEXT
hiredate	TIMESTAMP_NTZ
annual_income	NUMBER
salary_range	TEXT
education	TEXT
workschedule_translation	TEXT
education_level	TEXT
jobrole	TEXT
status	NUMBER

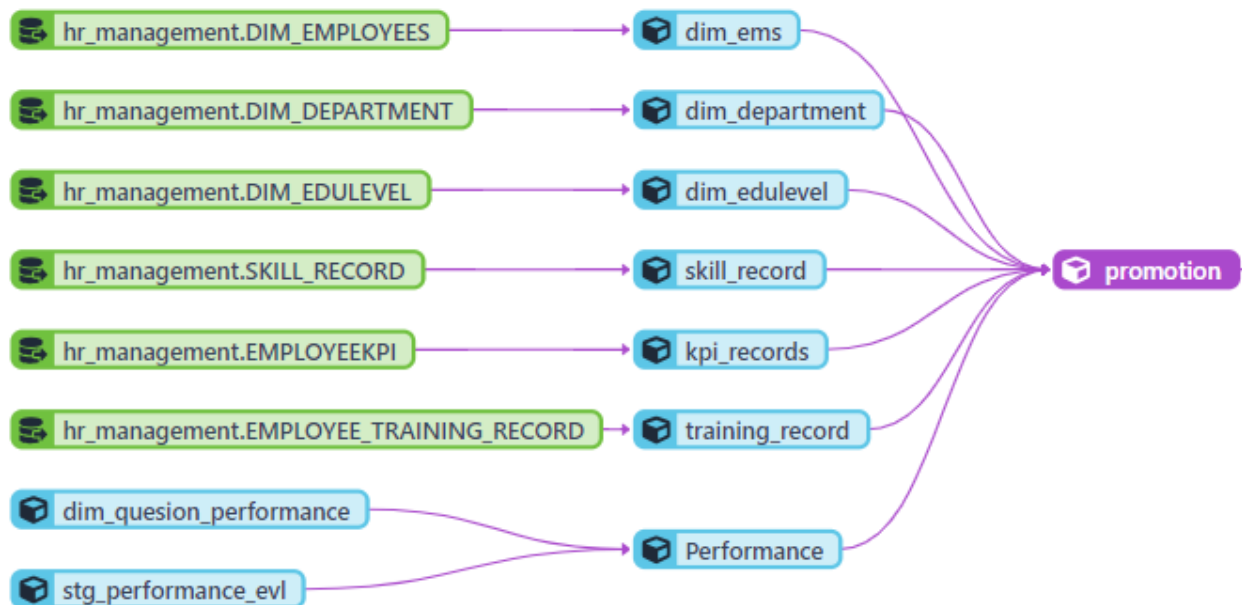
Bảng WORKFORCE

Với bảng OLE, thu thập dữ liệu từ số lần nghỉ của nhân viên sau các tháng, KPI theo tháng, loại công việc (full time-part time để tính số giờ làm việc tối thiểu), sau khi có đủ thông tin tính toán công thức chỉ số Overall Labor Effectiveness.

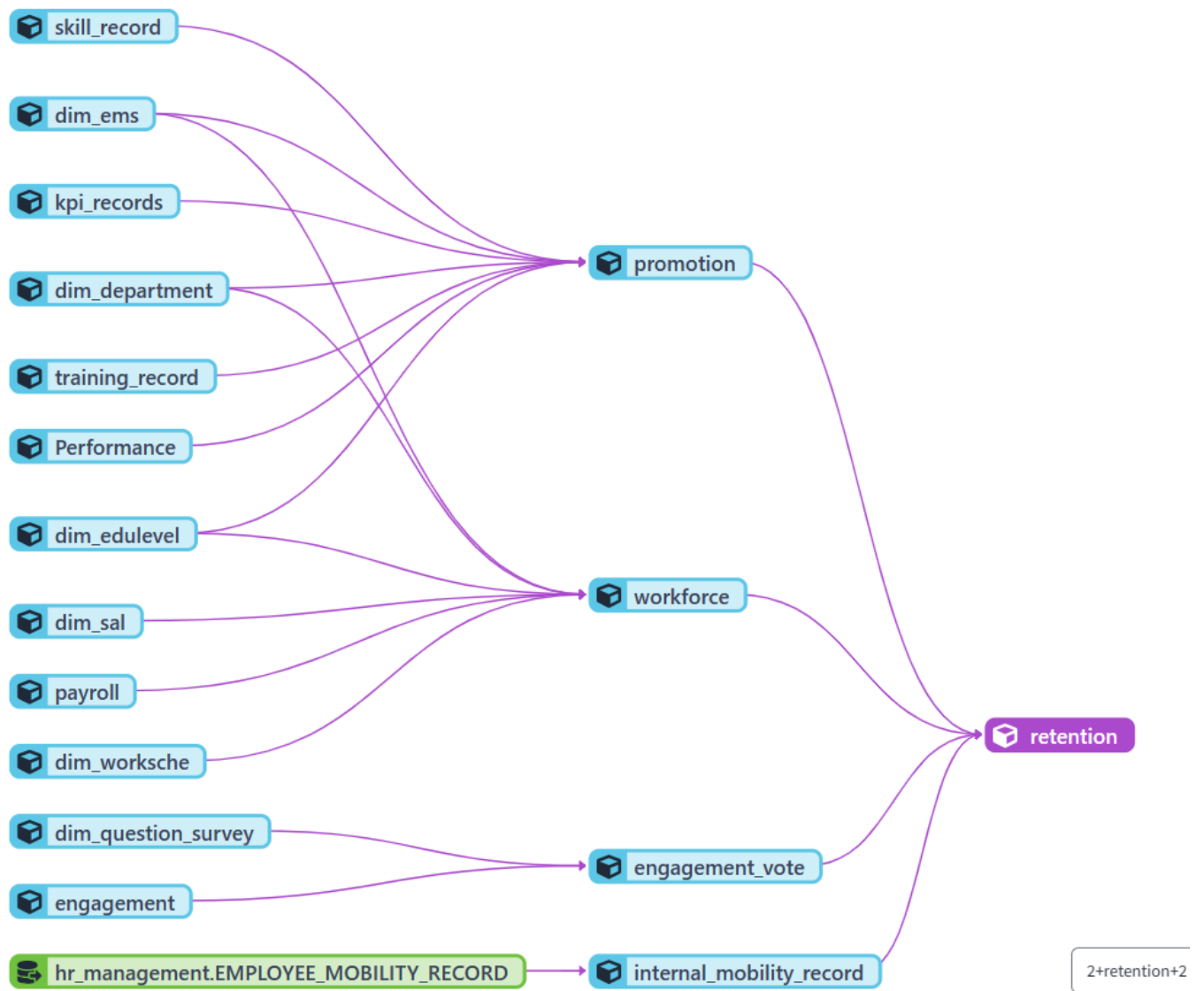


Linear Graph cho chỉ số OLE -Dashboard Performance

Bảng Promotion bao gồm các feature cần thiết cho việc dự đoán khả năng thăng tiến của nhân viên bao gồm thông tin cơ bản như phòng ban, trình độ học vấn, đến kpi hàng tháng của nhân viên, số lần tham gia training, và các bảng Performance cung cấp thông tin đánh giá của quản lý đối với nhân viên

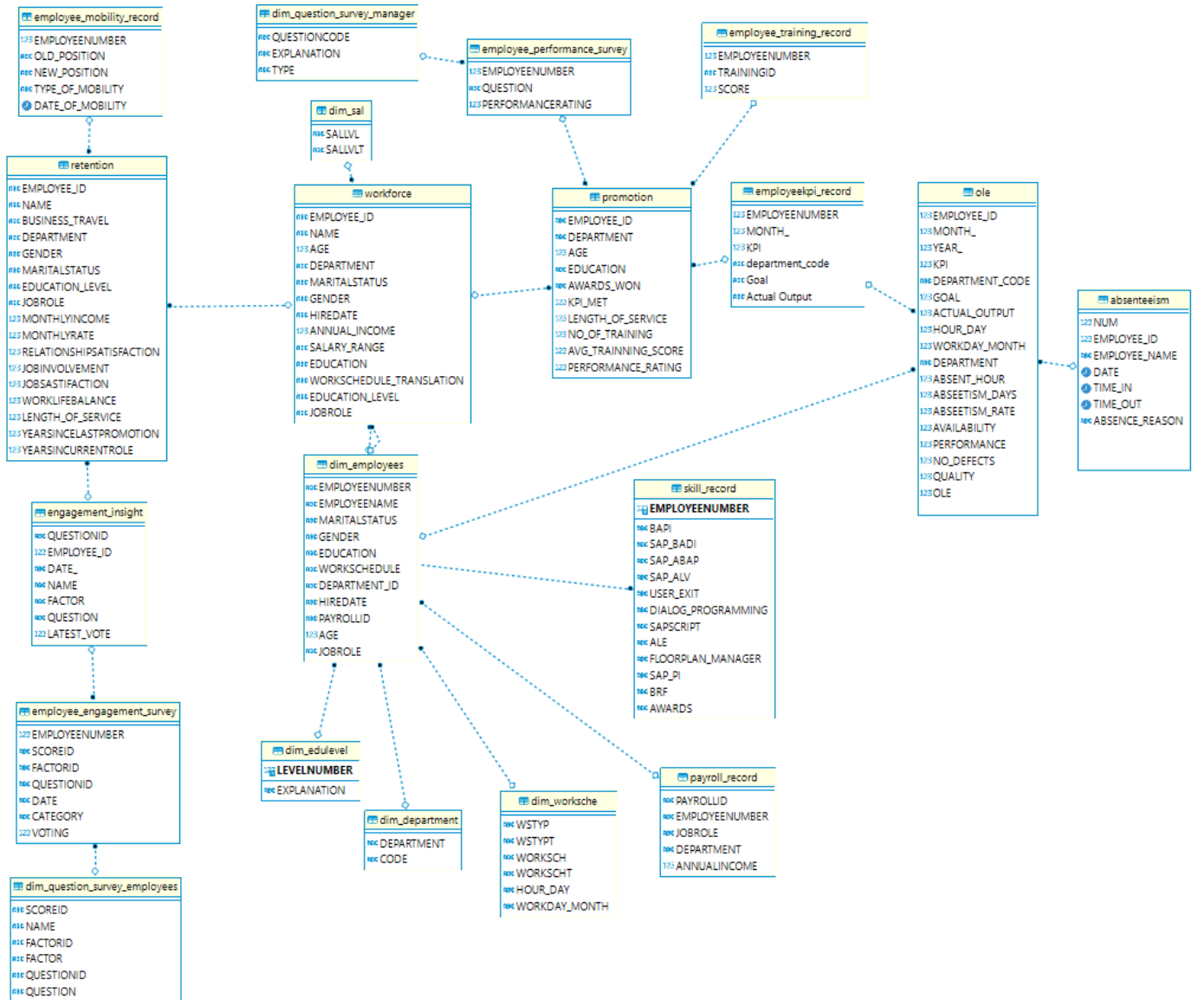


Linear Graph cho model data cần để dự đoán khả năng thăng tiến



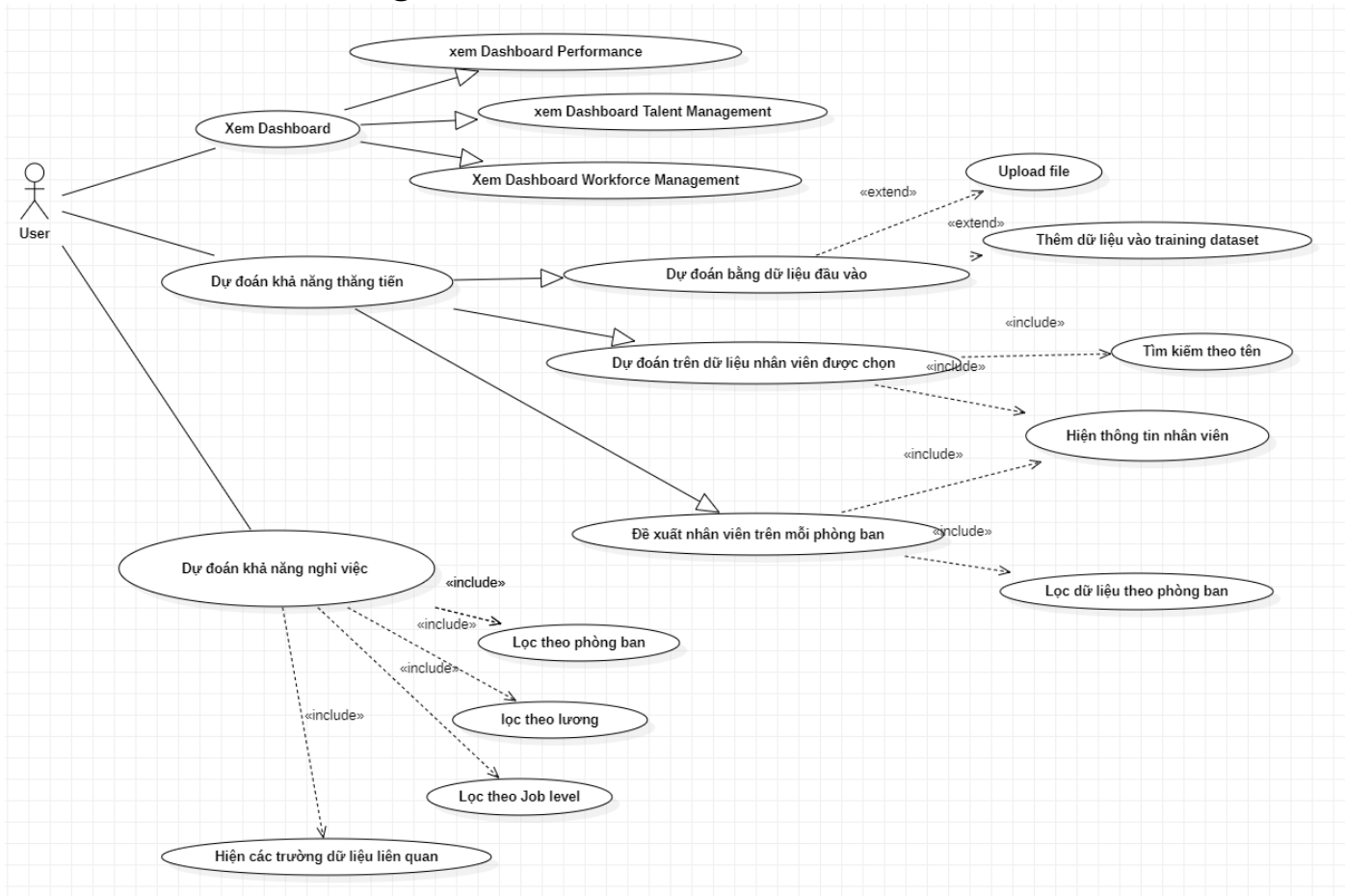
RETENTION cung cấp các feature cần thiết để dự đoán khả năng nghỉ việc

1.3. Data Model



2. Phân tích hệ thống

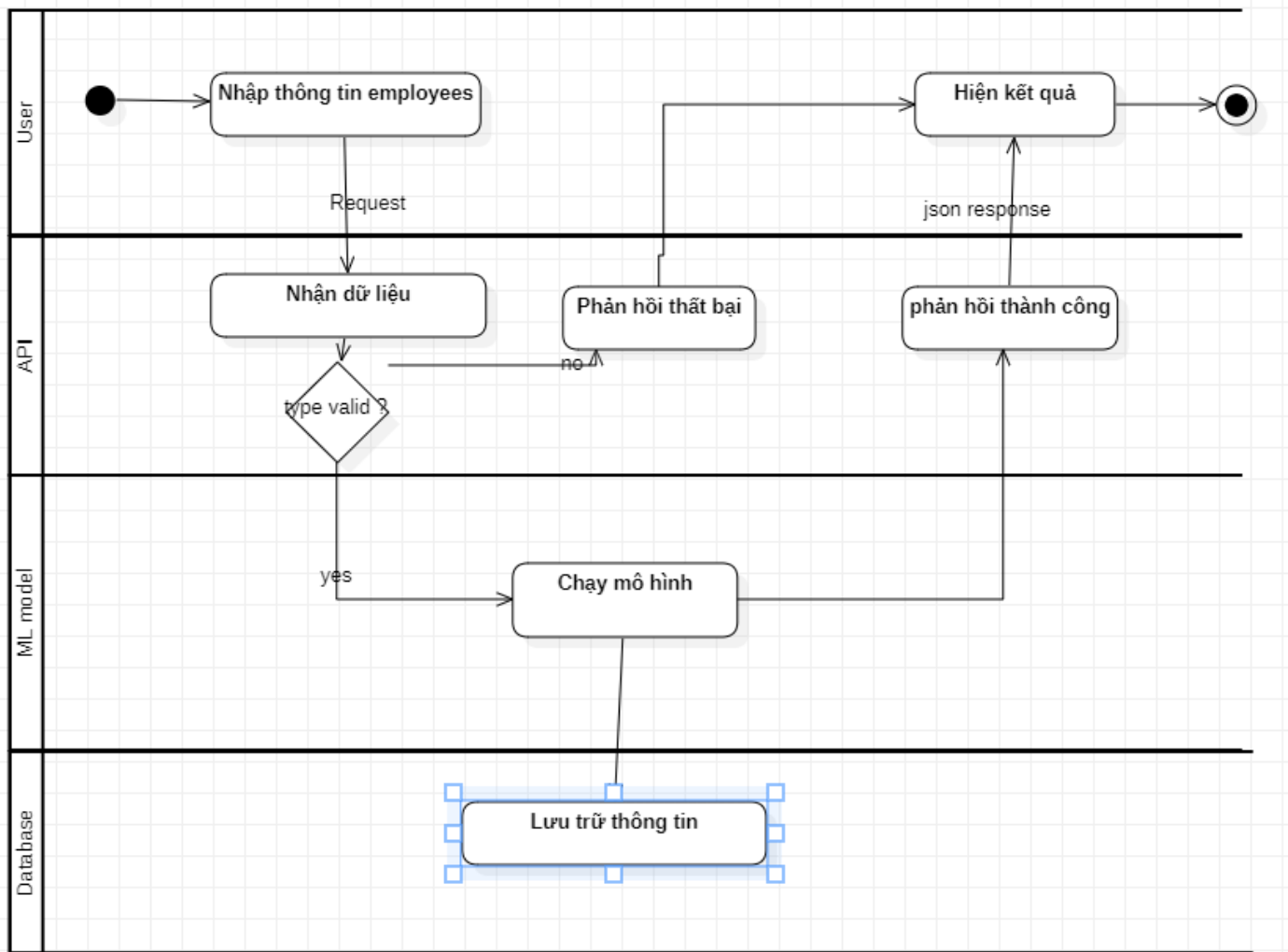
2.1. Use Case Diagram



Thực hiện 3 loại chức năng chính là:

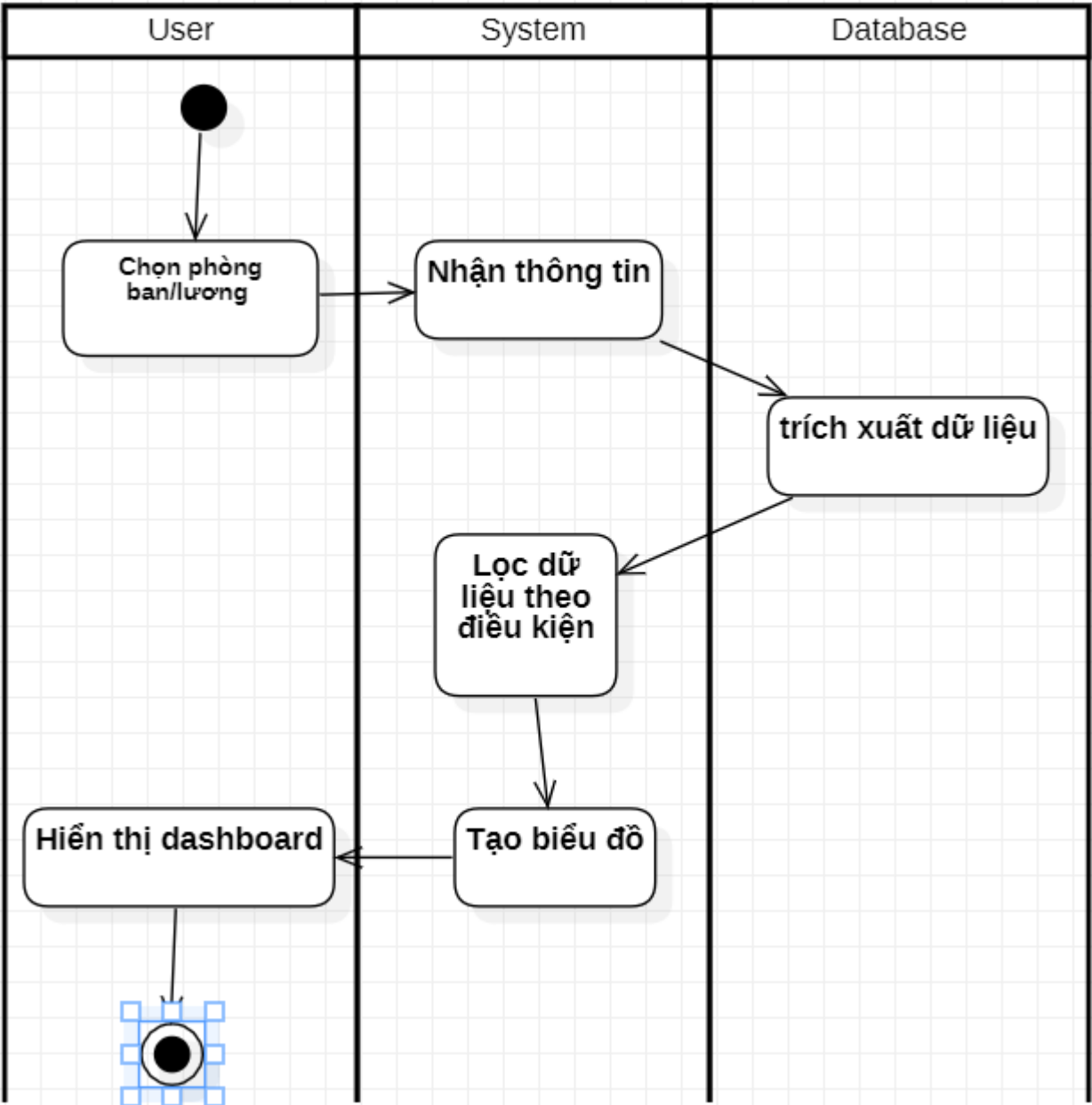
1. Chạy Dashboard cập nhập real-time theo đường đi của dữ liệu qua Snowflake and DBT. Sẽ có 3 Dashboard:
 - Workforce Management
 - Talent Management
 - Performance
2. Dự đoán nhân viên có tiềm năng tốt và có thể sẽ thăng cấp, đồng thời thể hiện các feature liên quan tới kết quả dự đoán
3. Dự đoán khả năng sẽ nghỉ việc của nhân viên, chọn lọc theo một số thông tin như phòng ban, lương. Đồng thời thể hiện các feature liên quan ảnh hưởng tới kết quả dự đoán. Nêu ra nhân viên có khả năng nghỉ việc trên 80%

2.2. Activity Diagram

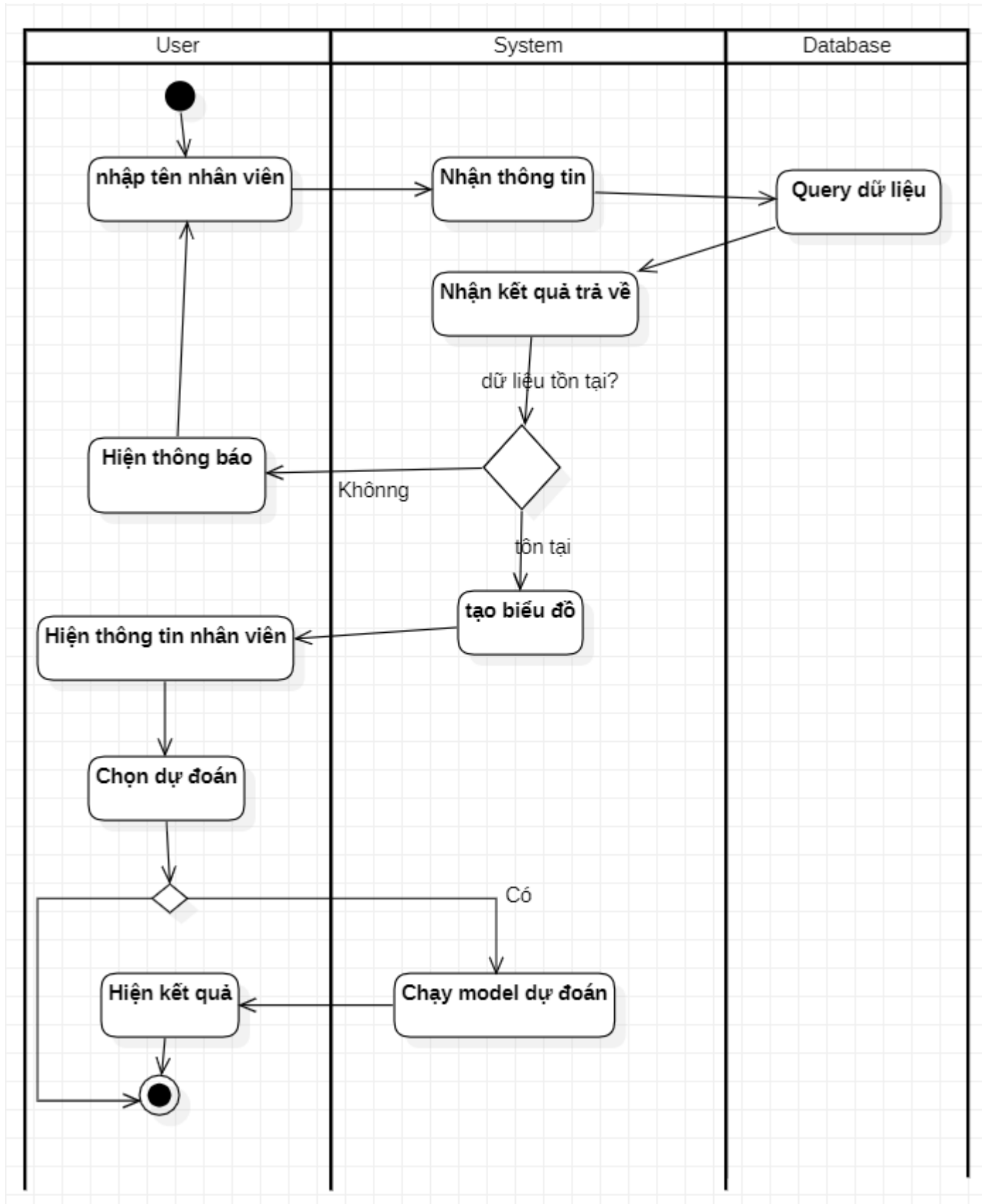


Activity chức năng dự đoán với model machine learning

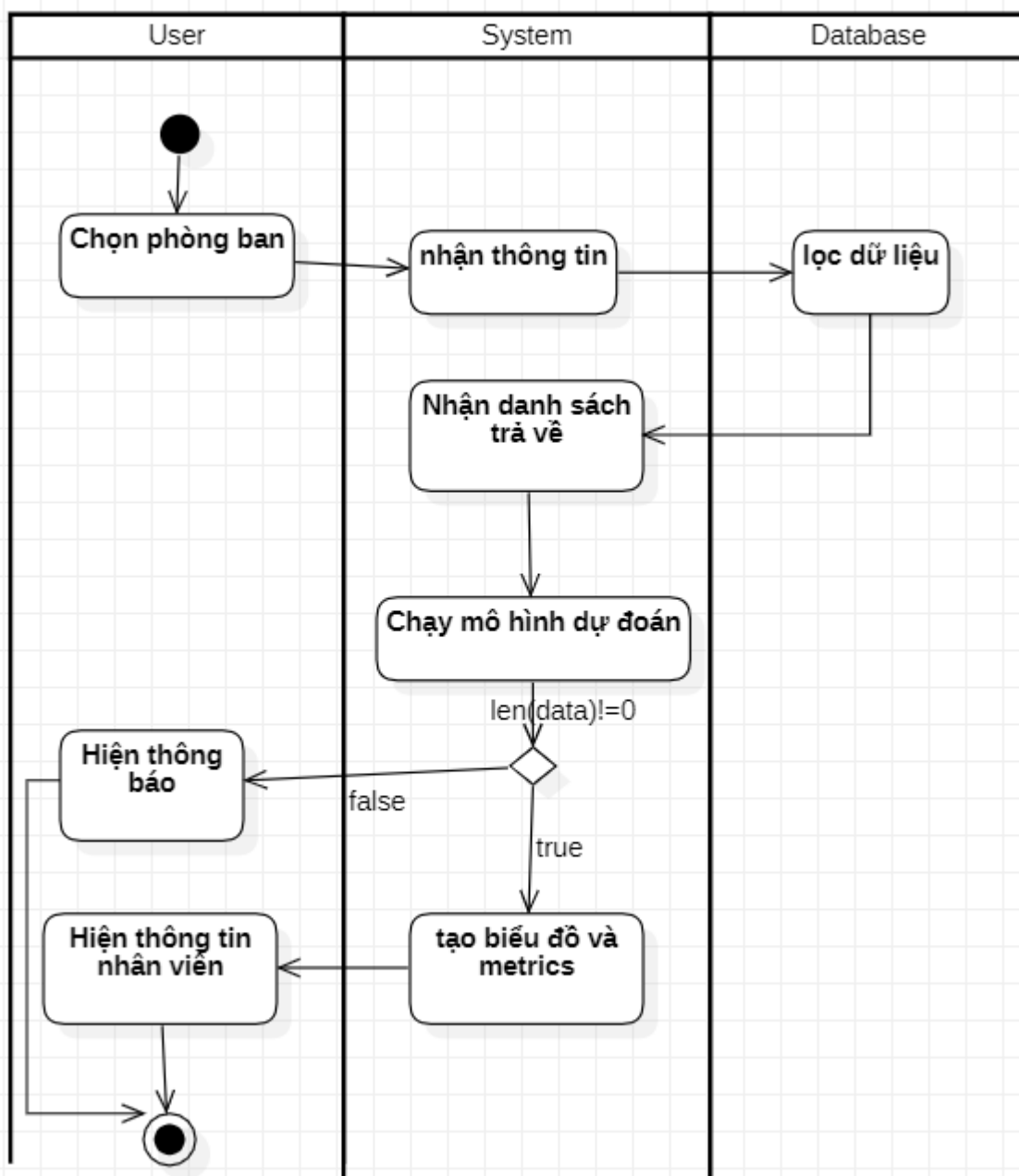
2.2.1 Xem Dashboard



2.2.2. Dự đoán khả năng thăng tiến

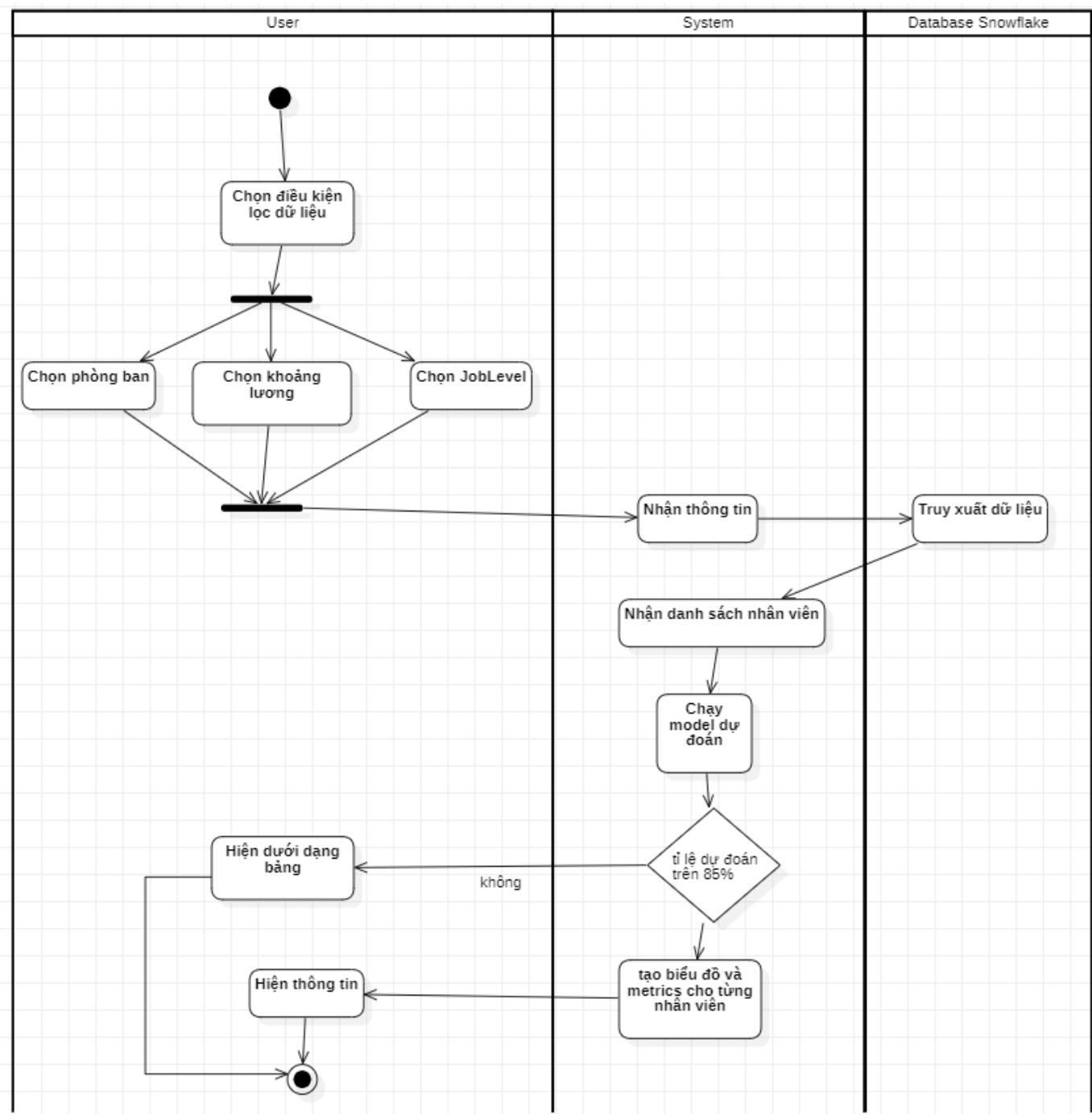


Activity cho chức năng dự đoán trên dữ liệu nhân viên được chọn



Dự đoán và đưa đề xuất trên mỗi phòng ban

2.2.3. Dự đoán khả năng nghỉ việc

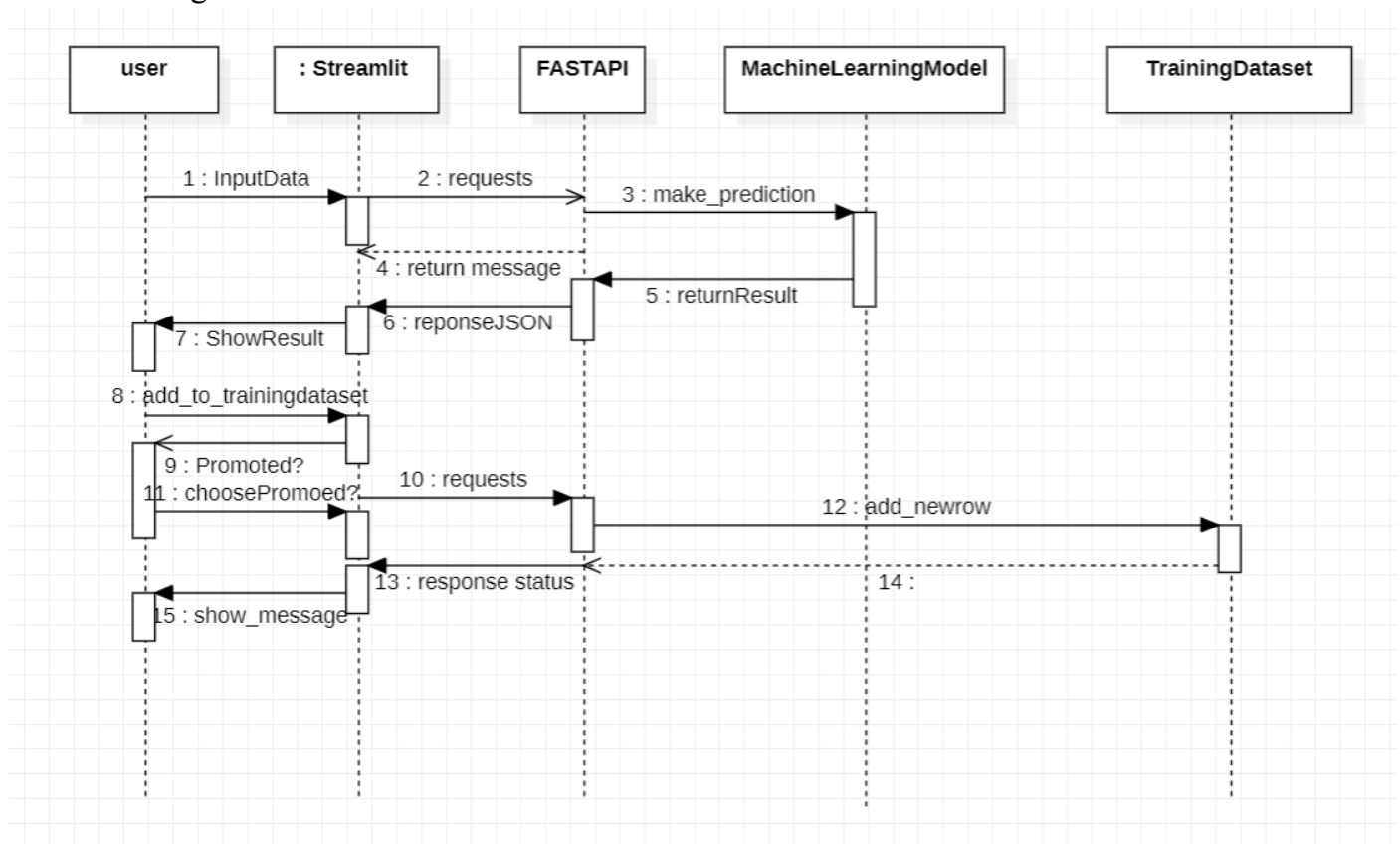


2.3 Sequence Diagram

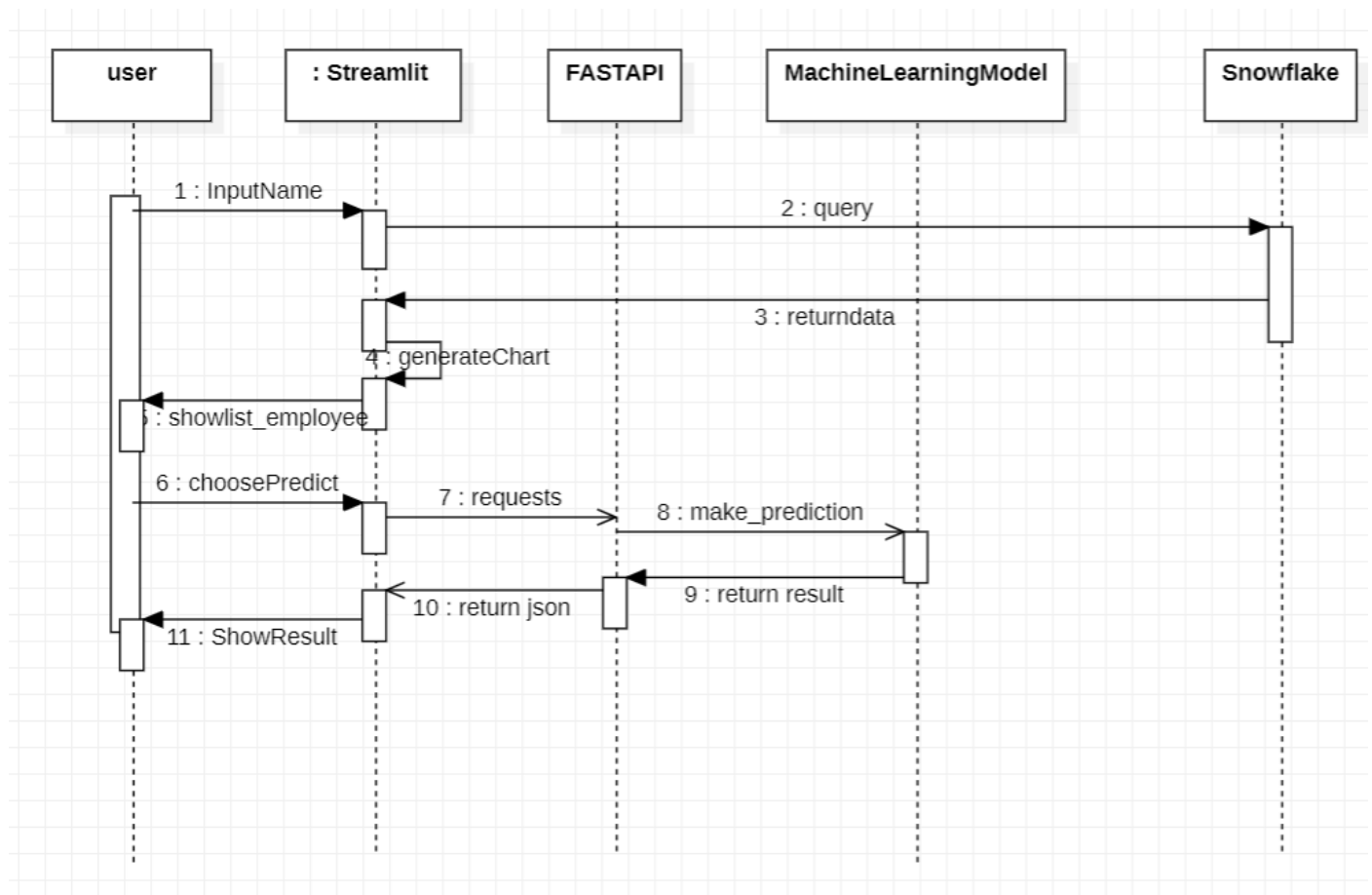
2.2.1. Dự đoán khả năng thăng tiến

Dự đoán trên các giá trị nhập vào từ người dùng

1. Người dùng nhập vào các giá trị cần theo các features dùng để train model
2. Streamlit gửi requests với FastAPI, từ đây giúp trích xuất dữ liệu và tương tác với model
3. Sau khi chạy model, gửi trả kết quả về interface với api, streamlit sẽ chuyển đổi file về dưới dạng dataframe và hiển thị kết quả dự đoán
4. Sau khi nhập các dữ liệu theo form, người dùng sẽ có lựa chọn có muốn thêm dữ liệu này vào file training data hay không, sau khi nhấn thêm, dòng giá trị mới sẽ được cập nhập vào file training dataset

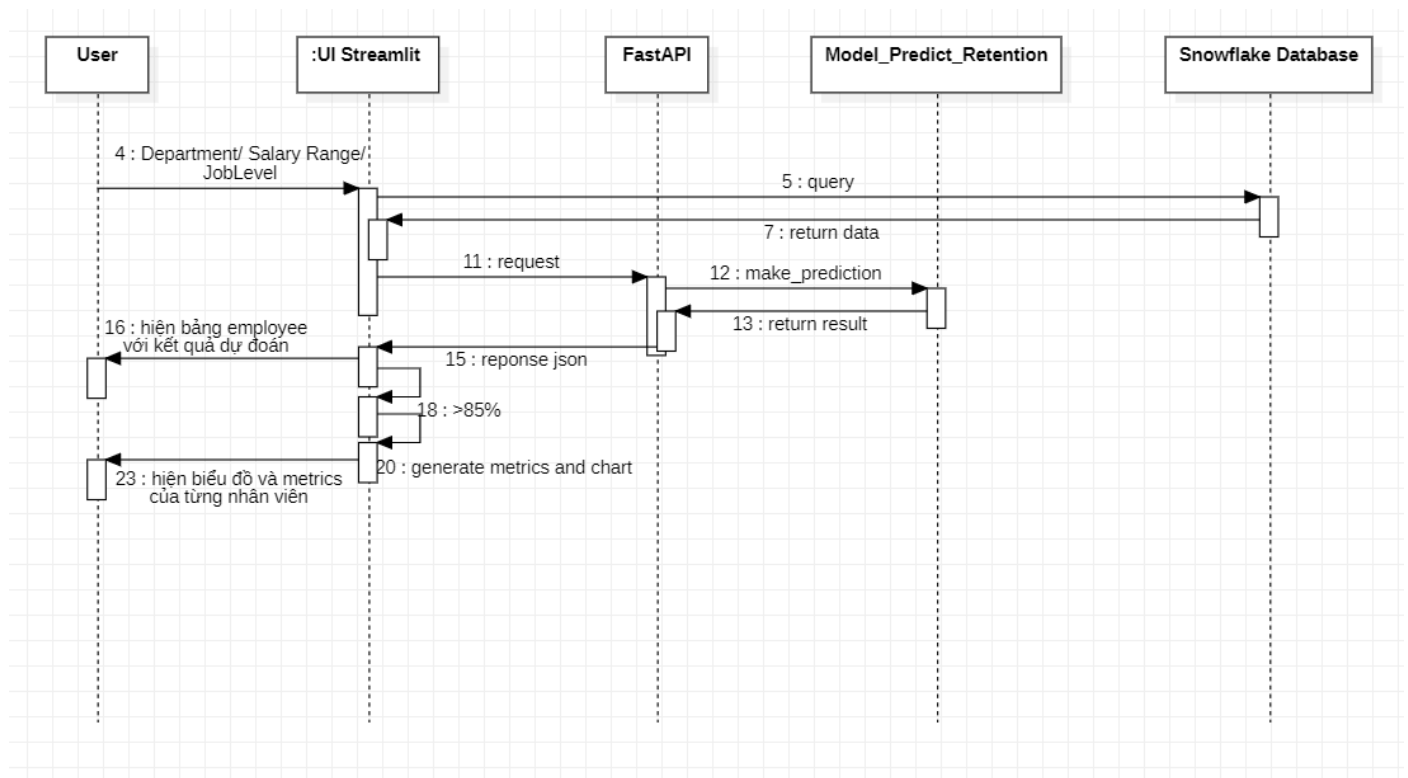


Dự đoán theo giá trị được nhập vào



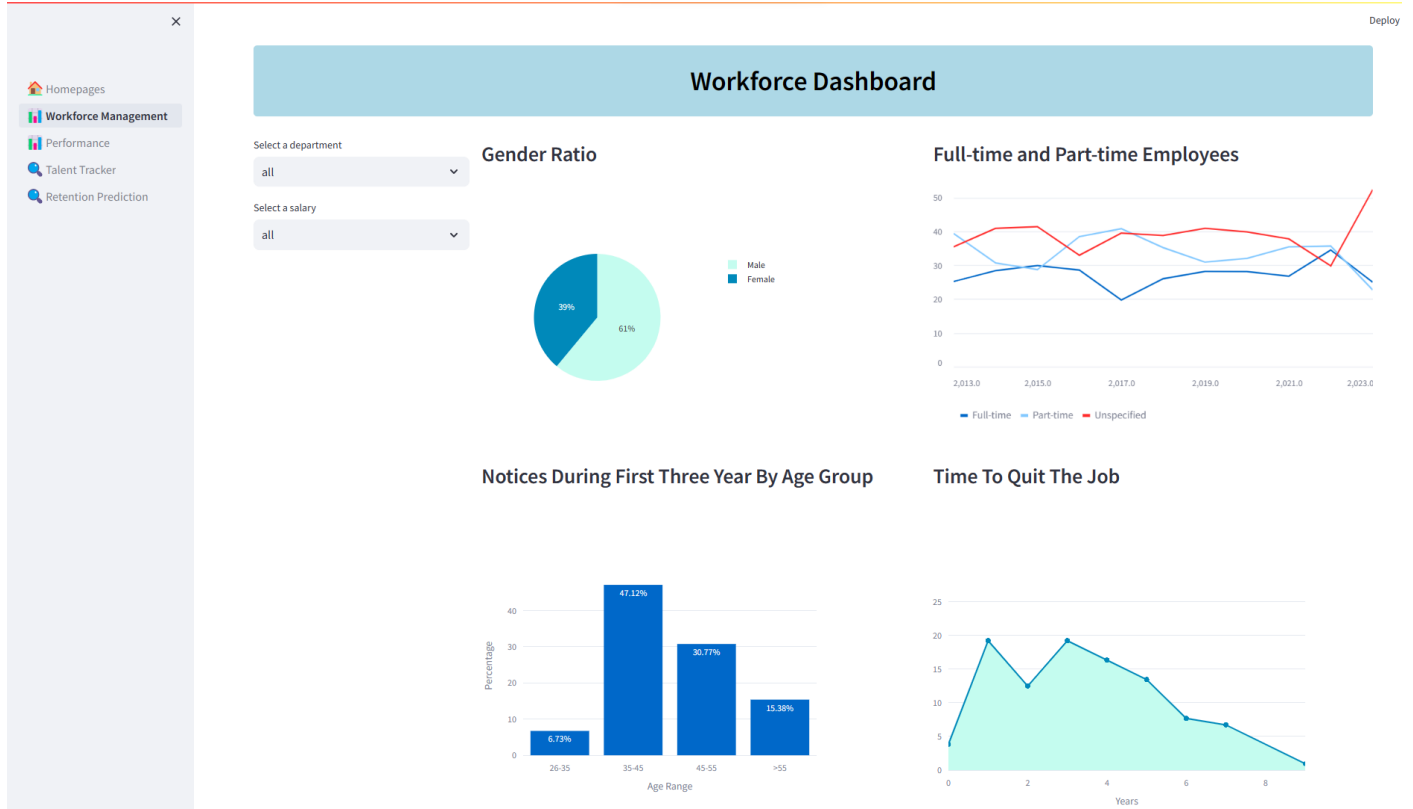
Dự đoán trên danh sách nhân viên chọn

2.2.2. Dự đoán khả năng nghỉ việc

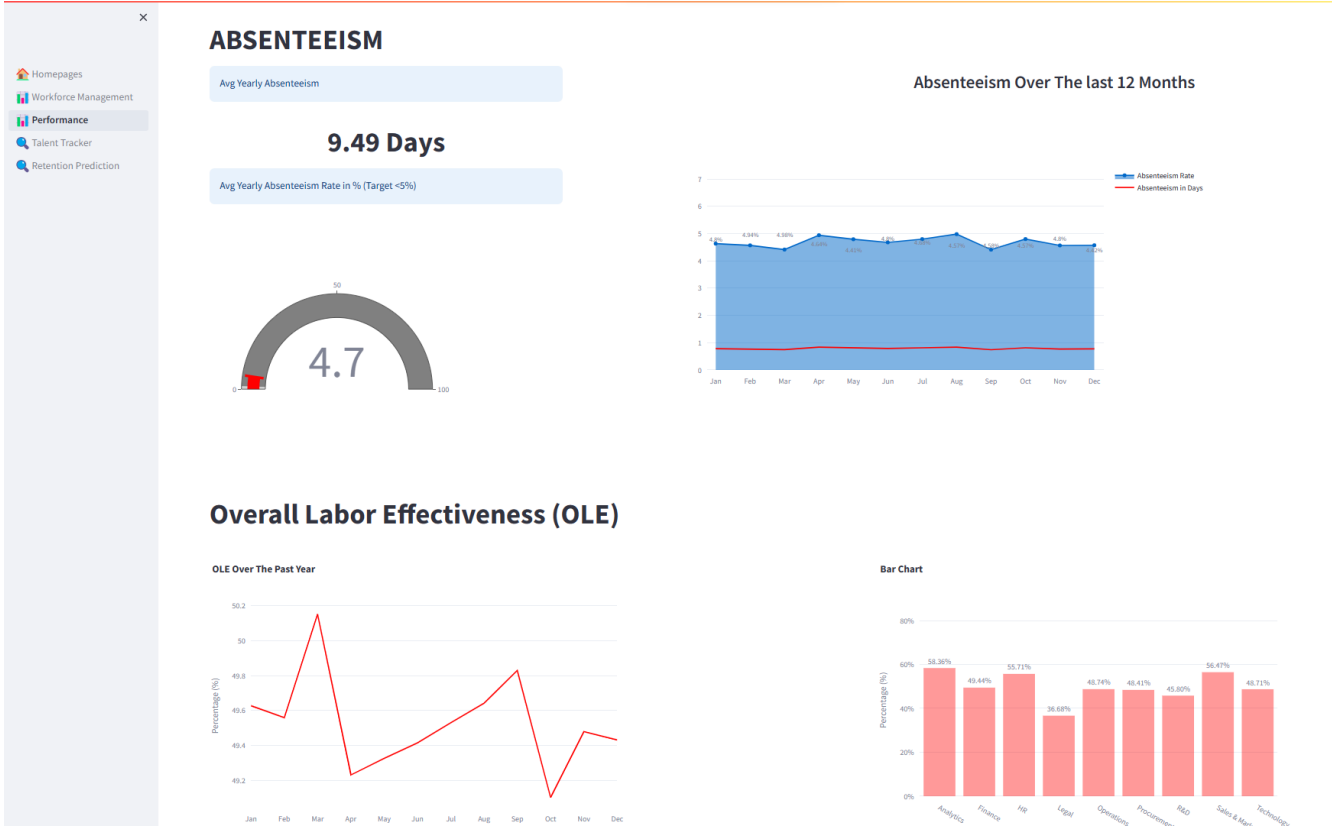


3 .Giao diện và Dashboard

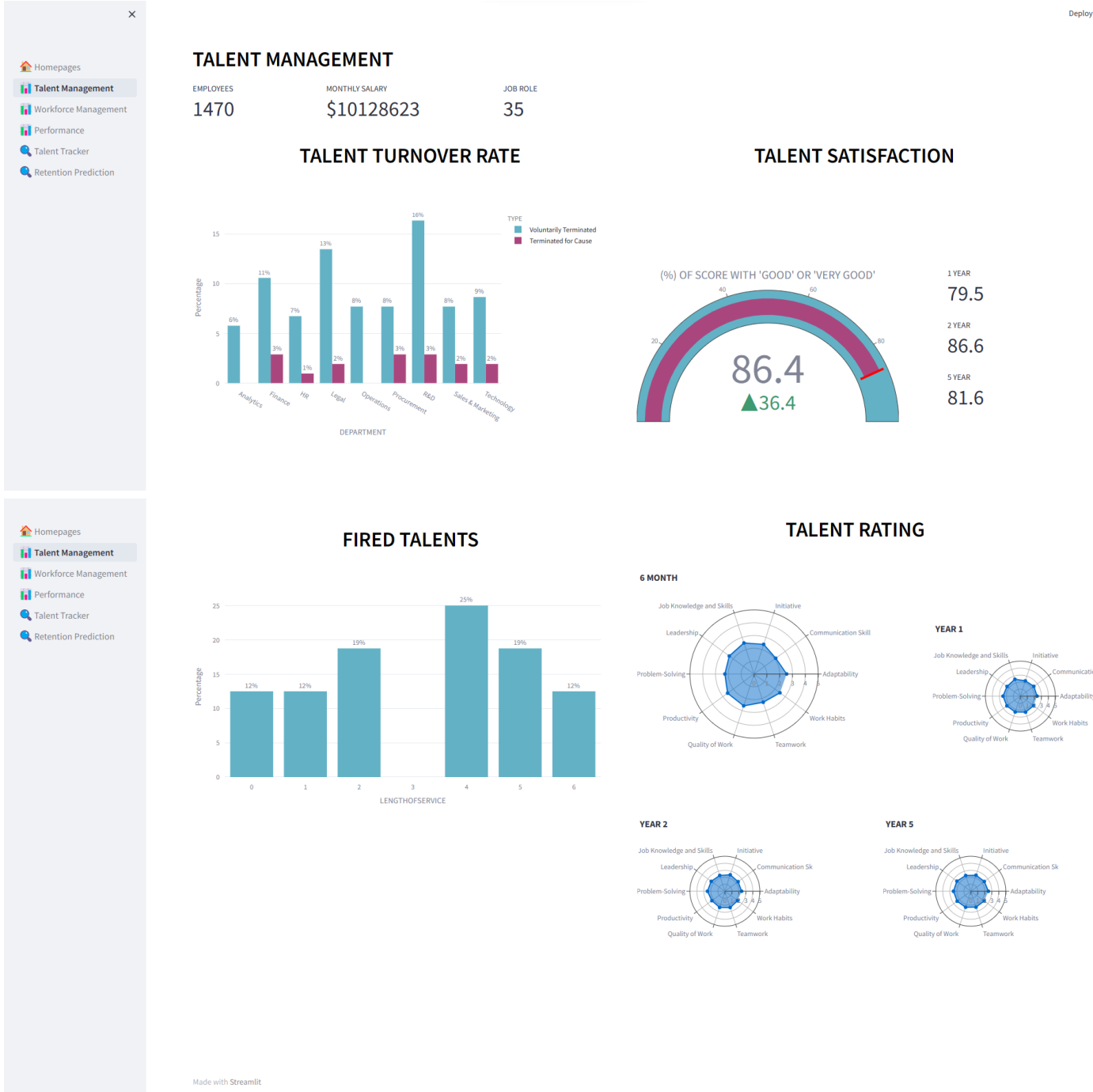
3.1. Dashboard Workforce Management



3.2. Dashboard Performance



3.4. Dashboard Talent Management



3.5. Dự đoán khả năng thăng chức với dữ liệu được nhập từ người dùng

×

Homepages

Workforce Management

Performance

Prediction

Make Prediction

Prediction - Potential Employee for Promotion

Using XGBoost, make predictions for which employees have potential to consider for promotion or praising them for their hard work

[Try Prediction](#) [Search for Employee](#) [Recommend by Department](#) [About the model](#)

☐ Upload Your Dataset

Region

region_7

Department

Sales & Marketing

Your highest Education

Doctor

Gender

☒ Female

☐ Male

Performance Rating

1

KPI met > 80%

☒ Yes

☐ No

Award

☒ Yes

☐ No

Recruitment Channel

☒ sourcing

☒ yes

☐ No

Award

☒ Yes

☐ No

Recruitment Channel

☒ sourcing

☐ other

☐ referred

Select Age

17

1765

Average Training Score

0

0100

Number of trainings

0.00

- +

Lenght of Service

0.00

- +

Predict

☐ Add to training dataset

Made with Streamlit

3.6. Thêm giá trị mới cho tranining dataset

☒ Add to training dataset

Promo?

☐ is_promo

☒ not_promo

Wanna add new row

New row has been added to training dataset!

3.7. Dự đoán trên danh sách nhân viên được chọn

×

Homepages

Workforce Management

Performance

Prediction

Make Prediction

Try Prediction

Search for Employee

Recommend by Department

About the model

Search by Name

Select employess

Tara Allison ×

Karen Salinas ×

Juan Thompson ×

Robert Harding ×

Stephanie Bryant ×

Courtney Scott ×

⌵

	employeenumber	employeeenname	maritalstatus	gender	education	workschedule	department_i
3	1860	Tara Allison	Single	0	3	J	R&D
11	1689	Karen Salinas	Married	0	4	H	S&M
18	1683	Juan Thompson	Married	0	4	P	PRO
20	421	Stephanie Bryant	Divorced	1	2	T	OPE
24	749	Robert Harding	Single	0	3	S	OPE
37	1779	Courtney Scott	Married	1	1	Q	LEG

TARA ALLISON

KAREN SALINAS

JUAN THOMPSON

STEPHANIE BRYANT

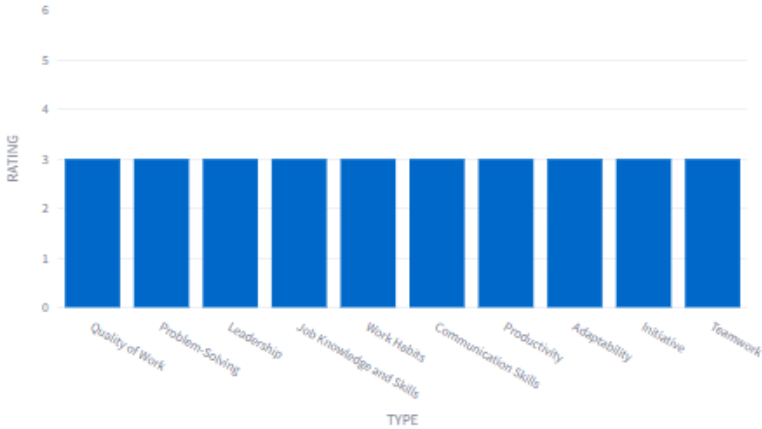
ROBERT HARDING

COURTNEY SCOTT

20	421	Stephanie Bryant	Divorced	1	2	T	OPE
24	749	Robert Harding	Single	0	3	S	OPE
37	1779	Courtney Scott	Married	1	1	Q	LEG

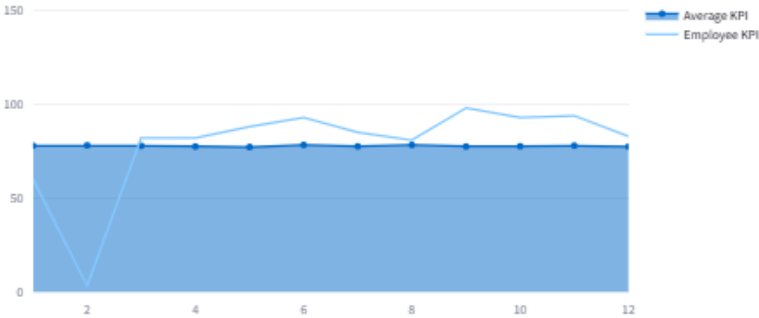
TARA ALLISON

Performance Rating from Manager



	employee_id	department	region	education	gender	recruitment_channel	no_of_training
3	1,860	R&D	region_7	Bachelor	1	sourcing	

KPI in 2023 by Month compare with the average



Predict

Request successful!

Predicted Status: This Employee needs more time to gain performance

KAREN SALINAS

JUAN THOMPSON

3.8 Dự đoán và đề xuất nhân viên trên mỗi phòng ban

×

Homepages

Workforce Management

Performance

Prediction

Make Prediction

Prediction - Potential Employee for Promotion

Using XGBoost, make predictions for which employees have potential to consider for promotion or praising them for their hard work

Try Prediction

Search for Employee

Recommend by Department

About the model

Potential Employees

Select a department

Technology

Technology

Finance

R&D

Analytics

Procurement

Sales & Marketing

Legal

Operations

×

Homepages

Workforce Management

Performance

Talent Tracker

Retention Prediction

Make Prediction

Potential Employees

Select a department

Technology

Predict

Request successful!

	employee_id	department	region	education	gender	no_of_trainings	age	previous_year_ratr
14	252	8	6	3	1	3	30	
30	286	8	45	2	0	2	40	
43	1,107	8	4	4	1	4	26	
48	52	8	4	3	0	6	38	
49	1,001	8	20	2	1	4	28	
57	804	8	45	1	0	3	51	
67	1,304	8	45	2	1	2	49	
80	96	8	3	3	0	3	32	
90	1,415	8	20	4	0	4	34	
94	1,661	8	20	2	1	2	46	
96	512	8	45	3	0	2	31	

Find 5 potential employees

379

Deploy

×

Homepages

Workforce Management

Performance

Talent Tracker

Retention Prediction


Make Prediction

Deploy

1 of 1 predicted employees

379

252



Request successful!

Nicole Silva

Postion: Project Manager

Age 38

Education: Bachelor

OLE Yearly Average %

49.0

Number of Trainings

3

Previous Year Rating

3.0

Salary Range

\$60,000 - \$69,999

Average Training Score

91.67

Absenteeism in days

18.0

Skills

	BAPI	SAP_BADI	SAP_ABAP	SAP_ALV	USER_EXIT	DIALOG_PROGRAMMING
0	good communications	freshers	management	Fresher	Graduate	events

1 2 3 4

3.9. Dự đoán nhân viên có xác suất nghỉ việc cao

×

Homepages

Workforce Management

Performance

Talent Tracker

Retention Prediction

Select a department
All

Select a salary range:
All

Choose Job Level
All

All

2

1

3

4

5

Deploy

List of all Employees have predicted result as likely to quit

	AGE	BUSINESS TRAVEL	DAILY RATE	DEPARTMENT	DISTANCE FROM HOME	EDUCATION	EDUCATION FIELD	EMPLOYEE NUMBER	ENVIRONMENT SATISFACTION	GENDER	HOURLY RATE	JOB INVOLVEMENT	JOB LEVEL
0	41	Travel_Rarely	1,102	Sales	1	2	Life Sciences	1	2	Female	94	3	2
2	37	Travel_Rarely	1,373	Research & Development	2	2	Other	4	4	Male	92	2	1
14	28	Travel_Rarely	103	Research & Development	24	3	Life Sciences	19	3	Male	50	2	1
21	36	Travel_Rarely	1,218	Sales	9	4	Life Sciences	27	3	Male	82	2	1
26	32	Travel_Frequently	1,125	Research & Development	16	1	Life Sciences	33	2	Female	72	1	1
36	50	Travel_Rarely	869	Sales	3	2	Marketing	47	1	Male	86	2	1
42	26	Travel_Rarely	1,357	Research & Development	25	3	Life Sciences	55	1	Male	48	1	1
50	48	Travel_Rarely	626	Research & Development	1	2	Life Sciences	64	1	Male	98	2	3
51	28	Travel_Rarely	1,434	Research & Development	5	4	Technical Degree	65	3	Male	50	3	1
54	26	Travel_Rarely	1,443	Sales	23	3	Marketing	72	3	Female	47	2	2

Found 15 employees likely to quit with probability of more than 85%

31

40

94

78

39

71

92

×

Homepages

Workforce Management

Performance

Talent Tracker

Retention Prediction

Select a department

All

Select a salary range:

\$60000-\$79999

Choose Job Level

All

591	33	Travel_Rarely	118	Sales	16	3	Marketing	819	1	Female	69
601	40	Travel_Frequently	720	Research & Development	16	4	Medical	832	1	Male	51

Found 1 employees likely to quit with probability of more than 85%

16

Amanda Mejia

Probability to quit (%)

87.51

Monthly Income (\$)

5765

Age

29

Total working years

7

Job Role-Level: 2

Sales Executive

Bar chart showing satisfaction metrics for Amanda Mejia. The Y-axis ranges from 0 to 3. The X-axis categories are RELATIONSHIPSATISFACT, JOBSATISFACT, JOBINVOLEMENT, ENVIRONMENTSATISFACT, PERFORMANCECERATING, and WORKLIFEBALANCE. Blue bars represent the Mean Value, and a light blue line with markers represents the Row Value.

Metric	Mean Value	Row Value
RELATIONSHIPSATISFACT	2.8	1.0
JOBSATISFACT	2.8	2.0
JOBINVOLEMENT	2.8	3.0
ENVIRONMENTSATISFACT	2.8	1.0
PERFORMANCECERATING	3.2	3.0
WORKLIFEBALANCE	2.8	1.0

Salary Comparison

Bar chart comparing Employee Salary and Average Salary. The Y-axis ranges from 0 to 6000. Employee Salary is approximately 5800, and Average Salary is approximately 6200.

Category	Salary
Employee Salary	5800
Average Salary	6200

III. Tổng kết

1. Kiến thức tìm hiểu

1.1. Streamlit

Streamlit là một công cụ giúp xây dựng các ứng dụng web tương tác mới được ra mắt vào năm cuối năm 2019. Nó nhanh chóng ghi điểm bởi tính linh hoạt, và dễ dàng tùy chỉnh data có độ phức tạp cao khi cho phép tạo ra các giao diện người dùng mạnh mẽ và chuyển đổi các notebook Jupyter hoặc script Python thành các ứng dụng tương tác. Streamlit rất phù hợp cho việc xây dựng các ứng dụng phân tích dữ liệu và trực quan hóa.

Vì là website thiên về làm việc với dữ liệu, streamlit tối giản hóa giao diện không yêu cầu về html hay css, streamlit có thể tự động căn chỉnh, nhưng vẫn có thể tích hợp các câu lệnh html và css nếu muốn chỉnh sửa nâng cao hơn. Đối với doanh nghiệp đây là một yếu tố có lợi về chi phí thiết kế và lắp đặt,

1.2. Snowflake và Data Build Tool

Để thêm vào yếu tố real time cho dashboard, dữ liệu yêu cầu phải được cập nhật thường xuyên, từ dữ liệu ở hệ thống chính, được đưa lên s3 bucket, tại đây có thể dễ dàng tập hợp các file ở nhiều nguồn khác nhau, sau đó cập nhật vào Snowflake. Snowflake là một kho dữ liệu đám mây cung cấp khả năng lưu trữ và xử lý dữ liệu lớn, trong khi dbt là một công cụ giúp bạn xây dựng các mô hình dữ liệu trong kho dữ liệu của . Sự kết hợp của Snowflake và dbt cho phép xây dựng, kiểm tra và triển khai các mô hình dữ liệu một cách hiệu quả.

1.3. Về lĩnh vực quản trị nhân sự

Công thức để tính Hiệu quả Lao động Tổng thể (OLE) trong ngành Nhân sự là:

$$OLE=Availability\times Performance\times Quality$$

Availability: Giả sử một nhân viên làm việc 8 giờ mỗi ngày, nhưng chỉ dành 7 giờ để làm việc thực sự (còn lại là thời gian nghỉ giải lao, họp...). Vậy thì Availability của nhân viên này là $7/8 = 87.5\%$.

Performance (Hiệu suất): Giả sử trong 7 giờ làm việc, nhân viên hoàn thành được 70% công việc được giao. Vậy thì Performance của nhân viên này là 70%.

Quality (Chất lượng): Giả sử trong số công việc mà nhân viên hoàn thành, có 95% công việc đạt chất lượng yêu cầu. Vậy thì Quality của nhân viên này là 95%.

Thay các giá trị vào công thức, ta được:

$$OLE=0.875\times 0.7\times 0.95=0.5803125.$$

1.4. Kết hợp website với các yếu tố khoa học dữ liệu

Hiểu về Khoa học dữ liệu (Data Science): Khoa học dữ liệu là một lĩnh vực rộng lớn bao gồm nhiều kỹ thuật và công cụ khác nhau để phân tích và hiểu dữ liệu. Nó bao gồm các phương pháp thống kê, học máy, và trực quan hóa dữ liệu.

Hiểu về Học máy (Machine Learning): Học máy là một phần quan trọng của Khoa học dữ liệu, nơi các thuật toán được sử dụng để học từ dữ liệu và đưa ra dự đoán hoặc quyết định mà không cần được lập trình cụ thể.

XGBoost: XGBoost là một thư viện tối ưu hóa cho việc tăng cường gradient, được thiết kế để có hiệu suất cao, linh hoạt. Nó thực hiện các thuật toán học máy trong khuôn khổ Gradient Boosting. XGBoost cung cấp một giải pháp tăng cường cây song song (còn được gọi là GBDT, GBM) giải quyết nhiều vấn đề khoa học dữ liệu một cách nhanh chóng và chính xác.

Hồi quy Logistic (Logistic Regression): Hồi quy logistic là một kỹ thuật thống kê được sử dụng để phân loại nhị phân. Nó có thể dự đoán kết quả của một biến phụ thuộc phân loại từ một hoặc nhiều biến dự đoán. Thay vì cung cấp giá trị chính xác là 0 hoặc 1, nó cung cấp các giá trị xác suất nằm giữa 0 và 1.

Quá trình xây dựng mô hình: Để xây dựng mô hình cho khoa học dữ liệu, cần tuân theo các bước sau: Hiểu vấn đề -> Trích xuất dữ liệu -> Làm sạch dữ liệu -> Phân tích dữ liệu -> Chọn tính năng -> Kết hợp thuật toán học máy -> Kiểm tra các mô hình -> Đưa mô hình vào sản xuất.

Huấn luyện Mô hình: Để huấn luyện mô hình, thường tuân theo các bước sau: Hiểu vấn đề -> Trích xuất dữ liệu -> Làm sạch dữ liệu -> Thực hiện phân tích dữ liệu khám phá -> Chọn tính năng -> Kết hợp thuật toán học máy -> Kiểm tra các mô hình -> Triển khai mô hình.

Tạo Mô hình: Để tạo mô hình cho khoa học dữ liệu, cần tuân theo các bước sau: Hiểu vấn đề -> Trích xuất dữ liệu -> Làm sạch dữ liệu -> Phân tích dữ liệu -> Chọn tính năng -> Kết hợp thuật toán học máy -> Kiểm tra các mô hình -> Đưa mô hình vào sản xuất.

Các công cụ và ngôn ngữ lập trình: Python và R là hai ngôn ngữ lập trình phổ biến nhất trong khoa học dữ liệu¹. Các thư viện như Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow, PyTorch, và XGBoost đều rất hữu ích¹³. SQL cũng rất quan trọng để làm việc với cơ sở dữ liệu.

Thống kê: Kiến thức về thống kê cơ bản như kiểm định giả thuyết, phân phối chuẩn, và các phép đo tương quan cũng rất quan trọng.

Trực quan hóa dữ liệu: Kỹ năng trực quan hóa dữ liệu bằng cách sử dụng các biểu đồ và đồ thị để hiển thị thông tin một cách trực quan là rất quan trọng

2. Hướng phát triển

Quản trị nhân sự là lĩnh vực lớn và phức tạp, nhưng do chưa được các doanh nghiệp thực sự chú ý và đầu tư như các phòng ban khác nên phần lớn doanh nghiệp vẫn chưa thể khai thác tốt về tiềm năng con người. Sự phát triển của công nghệ đang giúp sự tối ưu hóa đến từng lĩnh vực trong cuộc sống và đối với doanh nghiệp nói riêng. Hệ thống tạo tiền đề xây dựng website ứng dụng phạm vi rộng lớn hơn trong lĩnh vực Data Science, với web có thể đơn giản hóa một số bước như thao tác data trực tiếp trên web, thêm feature cho data training, thêm biểu đồ và khả năng chọn dữ liệu thích hợp. Từ việc ứng dụng khả năng dự đoán của machine learning model, có thể thêm các chức năng khác như đề xuất phát triển cho từng nhân viên, hoặc làm rõ các yếu tố quyết định kết quả dự đoán. Ngoài ra có thể tích hợp với hệ thống dùng để chấm công, để cập nhật trực tiếp dữ liệu thời gian ra vào của nhân viên.

3. Tài liệu tham khảo

7 Best HR Metrics Dashboard Examples 2023 -hr.university
(<https://hr.university/analytics/hr-metrics-dashboard/>)

Employee Attrition Prediction Using Machine Learning - Taught by:Ryan Ahmed, Adjunct Professor & AI Enthusiast- Coursera
(<https://www.coursera.org/learn/employee-attrition-prediction/>)

Sách: The Adoption and Effect of Artificial Intelligence on Human Resources Management
Tác giả: Pallavi Tyagi, Naveen Chilamkurti, Simon Grima, Kiran Sood, Balamurugan Balusamy,

Ibm HR Analytics Employee Attrition & Performance - Dataset

LINK GITHUB source code: [Lilili2214/DataScienceWebBased_humanrouscses \(github.com\)](https://github.com/Lilili2214/DataScienceWebBased_humanrouscses)
https://github.com/Lilili2214/DataScienceWebBased_humanrouscses

LINK GITHUB DBT: [Lilili2214/streamlit_application at DBT-model \(github.com\)](https://github.com/Lilili2214/streamlit_application_at_DBT-model)
https://github.com/Lilili2214/streamlit_application/tree/DBT-model

