**Dự án Phân tích Nguồn nhân lực**

**Tên dự án:** Nâng cao hiểu biết về Nguồn nhân lực thông qua Phân tích nâng cao

**Tóm tắt điều hành**

Mục tiêu của dự án là tận dụng phân tích dữ liệu và học máy để có được cái nhìn sâu sắc hơn về hoạt động nguồn nhân lực (HR) của chúng ta. Bằng cách phân tích bộ dữ liệu HR được cung cấp, chúng tôi đặt mục tiêu xác định các xu hướng và mẫu có thể giúp tối ưu hóa quản lý nhân lực, cải thiện sự hài lòng của nhân viên và nâng cao hiệu quả tổ chức chung. Dự án sẽ tập trung vào các chỉ số HR khác nhau như nhân khẩu học của nhân viên, điểm hiệu suất và chi tiết về bồi thường, với trọng tâm mạnh mẽ vào phân tích dự đoán.

**Mục tiêu và Mục tiêu**

**Mục tiêu:**

1. **Cải thiện Quản lý Nhân lực:** Nâng cao năng lực ra quyết định trong các thực hành HR để tối ưu hóa phân bổ và sử dụng nhân viên.
2. **Tăng cường Sự hài lòng của Nhân viên:** Xác định các yếu tố ảnh hưởng đến sự hài lòng và hiệu suất của nhân viên để thực hiện các chiến lược giữ chân nhân viên hiệu quả.
3. **Tối ưu hóa Chiến lược Bồi thường:** Phân tích sự phân bố lương và tiền thưởng để đảm bảo các thực hành bồi thường cạnh tranh và công bằng.
4. **Cái nhìn sâu sắc Dự đoán:** Phát triển các mô hình dự đoán để dự báo tình trạng nghỉ việc và hiệu suất của nhân viên nhằm chủ động giải quyết các vấn đề tiềm ẩn.

**Mục tiêu:**

1. **Thực hiện Phân tích Dữ liệu Khám phá (EDA):** Hiểu bộ dữ liệu, xác định các xu hướng chính và phát hiện các mẫu ẩn.
2. **Tạo Cái nhìn sâu sắc Thị giác:** Tạo các hình ảnh trực quan toàn diện để minh họa các phát hiện và hỗ trợ ra quyết định.
3. **Xác định các Chỉ số chính:** Tập trung vào các chỉ số HR quan trọng như tỷ lệ luân chuyển nhân viên, điểm hiệu suất và chênh lệch bồi thường.
4. **Phát triển Mô hình Dự đoán:** Xây dựng và đánh giá các mô hình để dự đoán tình trạng nghỉ việc và hiệu suất của nhân viên.
5. **Cung cấp Khuyến nghị:** Đưa ra các khuyến nghị dựa trên dữ liệu để cải thiện quy trình và chính sách HR dựa trên kết quả phân tích và mô hình dự đoán.

**Yêu cầu dành cho Chuyên viên Phân tích Dữ liệu**

**Chuẩn bị Dữ liệu chung:**

1. **Chất lượng và Độ chính xác của Dữ liệu**
   * Có bất kỳ giá trị nào bị thiếu, không nhất quán hoặc ngoại lệ trong bộ dữ liệu không? Bạn sẽ xử lý chúng như thế nào?
   * Bạn sẽ thực hiện những bước nào để đảm bảo dữ liệu sạch và được chuẩn hóa?
2. **Kỹ thuật Tạo Tính năng**
   * Các tính năng mới nào có thể được tạo ra từ dữ liệu hiện có để tăng cường sức mạnh dự đoán của các mô hình (ví dụ: thâm niên, nhóm tuổi)?
   * Bạn sẽ chuyển đổi các biến loại thành định dạng số như thế nào (ví dụ: mã hóa một nóng)?

**Phân tích Dữ liệu Khám phá (EDA):**

1. **Thống kê Mô tả**
   * Thống kê tóm tắt (trung bình, trung vị, độ lệch chuẩn) cho các tính năng số là gì?
   * Bạn sẽ xử lý và giải thích phân phối của các tính năng khác nhau như thế nào?
2. **Hình ảnh trực quan**
   * Bạn sẽ tạo các loại hình ảnh trực quan nào để khám phá dữ liệu (ví dụ: biểu đồ cột, biểu đồ histogram, biểu đồ tròn, biểu đồ phân tán)?
   * Bạn sẽ trực quan hóa sự phân bố nhân viên theo bộ phận, vị trí, quốc gia và thành phố như thế nào?
   * Bạn sẽ hiển thị sự phân bố lương và tiền thưởng như thế nào?

**Mô hình Dự đoán:**

1. **Lựa chọn Mô hình**
   * Bạn sẽ xem xét các thuật toán học máy nào để dự đoán tình trạng nghỉ việc và hiệu suất của nhân viên? Tại sao?
   * Bạn sẽ biện minh cho sự lựa chọn thuật toán dựa trên đặc điểm của dữ liệu và vấn đề đang được giải quyết như thế nào?
2. **Huấn luyện Mô hình và Điều chỉnh Siêu tham số**
   * Bạn sẽ sử dụng phương pháp nào để huấn luyện các mô hình (ví dụ: phân chia đào tạo-kiểm tra, xác thực chéo)?
   * Bạn sẽ thực hiện điều chỉnh siêu tham số như thế nào (ví dụ: tìm kiếm lưới, tìm kiếm ngẫu nhiên)?
3. **Đánh giá Mô hình**
   * Bạn sẽ sử dụng các chỉ số nào để đánh giá hiệu suất của các mô hình dự đoán tình trạng nghỉ việc (ví dụ: độ chính xác, độ chính xác, độ thu hồi, F1-điểm, ROC-AUC)?
   * Bạn sẽ sử dụng các chỉ số nào để đánh giá hiệu suất của các mô hình dự đoán hiệu suất (ví dụ: MAE, MSE, RMSE, R-bình phương)?
   * Bạn sẽ đảm bảo các mô hình mạnh mẽ và có thể khái quát hóa như thế nào?
4. **Giải thích Mô hình**
   * Bạn sẽ giải thích kết quả của các mô hình như thế nào để hiểu các yếu tố dự đoán chính của tình trạng nghỉ việc và hiệu suất?
   * Bạn sẽ sử dụng kỹ thuật nào để giải thích các dự đoán của mô hình (ví dụ: tầm quan trọng của tính năng, giá trị SHAP, LIME)?

**Báo cáo và Tài liệu:**

1. **Báo cáo Kết quả Mô hình**
   * Bạn sẽ cấu trúc báo cáo ngắn gọn tóm tắt quy trình phát triển mô hình, kết quả đánh giá và các phát hiện chính như thế nào?
   * Bạn sẽ bao gồm các hình ảnh trực quan nào để minh họa hiệu suất của mô hình và các tính năng quan trọng?
2. **Nộp Mã nguồn**
   * Bạn sẽ tài liệu hóa mã nguồn như thế nào để đảm bảo nó sạch và có thể tái tạo?
   * Bạn sẽ thực hiện những bước nào để đảm bảo mã được chú thích và tổ chức tốt?
3. **Khuyến nghị**
   * Dựa trên kết quả của mô hình, bạn sẽ đưa ra những khuyến nghị khả thi nào để cải thiện việc giữ chân và hiệu suất của nhân viên?
   * Bạn sẽ đề xuất các chiến lược nào để tối ưu hóa bồi thường nhằm tối đa hóa sự hài lòng và năng suất của nhân viên?

**Tiêu chí Đánh giá (% Điểm)**

|  |  |
| --- | --- |
| Tiêu chí | Trọng lượng (%) |
| Làm sạch và Chuẩn bị Dữ liệu | 15 |
| Phân tích Dữ liệu Khám phá | 15 |
| Chất lượng Hình ảnh trực quan | 15 |
| Lựa chọn và Biện minh Mô hình | 10 |
| Huấn luyện Mô hình và Điều chỉnh Siêu tham số | 15 |
| Đánh giá và Giải thích Mô hình | 15 |
| Báo cáo và Tài liệu | 10 |
| Khuyến nghị và Khả thi | 5 |

**Mô tả Dữ liệu**

|  |  |
| --- | --- |
| Tên Cột | Mô tả |
| ID Nhân viên | Mã định danh duy nhất cho mỗi nhân viên. |
| Tên | Tên của nhân viên. |
| Giới tính | Giới tính của nhân viên (ví dụ: Nam, Nữ). |
| Tuổi | Tuổi của nhân viên. |
| Bộ phận | Bộ phận mà nhân viên làm việc (ví dụ: HR, Sales, IT). |
| Vị trí | Vị trí của nhân viên (ví dụ: Quản lý, Chuyên viên phân tích). |
| Lương | Lương của nhân viên. |
| Ngày vào làm | Ngày nhân viên tham gia công ty. |
| Điểm Hiệu suất | Điểm hiệu suất của nhân viên. |
| Tiền thưởng | Số tiền thưởng mà nhân viên nhận được. |
| Quốc gia | Quốc gia nơi nhân viên sinh sống. |
| Thành phố | Thành phố nơi nhân viên sinh sống. |

**Kết luận**

Bằng cách tuân theo các yêu cầu đã nêu và tập trung vào các mục tiêu và mục tiêu đã xác định, dự án này đặt mục tiêu cung cấp những hiểu biết khả thi về hoạt động HR của chúng ta. Phân tích dữ liệu và mô hình dự đoán sẽ giúp đưa ra các quyết định sáng suốt để cải thiện quản lý nhân lực, nâng cao sự hài lòng của nhân viên và đảm bảo các thực hành bồi thường công bằng.