Employee Attrition Analysis and Prediction

This project aims to provide insights into the factors influencing employee attrition and predict which employees are likely to leave the company.

Problem Statement:

Acme Corporation, a leading tech company, is facing a significant challenge with employee turnover. The HR department is concerned about the increasing rate of attrition, as it negatively impacts team dynamics, project continuity, and overall company morale. To address this issue, Acme Corporation wants to leverage data analytics and machine learning to understand the factors influencing employee turnover and predict which employees are likely to leave in the near future.

Dataset:

Acme Corporation has provided historical data on employee demographics, job satisfaction, work environment, performance metrics, and turnover status. This dataset spans the last five years and includes information on employees who have left the company and those who are still currently employed.

The dataset typically includes several features that provide insights into employee characteristics, job satisfaction, and performance. While the exact features may vary, here's a general list of common features you might find in such a dataset:

1. **Employee ID:** A unique identifier for each employee.
2. **Age:** The age of the employee.
3. **Attrition:** A binary variable indicating whether the employee has left the company (1) or is still employed (0).
4. **Business Travel:** The frequency and nature of business-related travel (e.g., "Travel_Rarely," "Travel_Frequently," "Non-Travel").

5. **Department:** The department to which the employee belongs (e.g., "Sales," "Research & Development," "Human Resources").
6. **Distance From Home:** The distance of the employee's residence from the workplace.
7. **Education:** The employee's level of education (e.g., "1: 'Below College'," "2: 'College'," "3: 'Bachelor'," "4: 'Master'," "5: 'Doctor').
8. **Education Field:** The field in which the employee's education lies (e.g., "Life Sciences," "Medical," "Marketing").
9. **Environment Satisfaction:** The level of satisfaction with the work environment on a scale.
10. **Gender:** The gender of the employee.
11. **Job Involvement:** The degree to which the employee is involved in their job.
12. **Job Level:** The level or rank of the employee's position.
13. **Job Role:** The specific role or title of the employee's job.
14. **Job Satisfaction:** The level of satisfaction with the job on a scale.
15. **Marital Status:** The marital status of the employee.
16. **Monthly Income:** The monthly salary of the employee.
17. **Num Companies Worked:** The number of companies the employee has worked for.
18. **Over Time:** Whether the employee works overtime or not.
19. **Performance Rating:** The performance rating of the employee.
20. **Relationship Satisfaction:** The level of satisfaction with relationships at the workplace.
21. **Stock Option Level:** The level of stock options provided to the employee.
22. **Total Working Years:** The total number of years the employee has been working.
23. **Training Times Last Year:** The number of training sessions the employee attended last year.
24. **Work-Life Balance:** The balance between work and personal life.
25. **Years At Company:** The number of years the employee has been with the current company.
26. **Years In Current Role:** The number of years the employee has been in their current role.
27. **Years Since Last Promotion:** The number of years since the last time the employee was promoted.
28. **Years With Current Manager:** The number of years the employee has been working under the current manager.

Please note that this is a general list, and the actual dataset might include additional features or variations. It's essential to explore the dataset thoroughly to understand the specifics of each feature and its relevance to the analysis.

Business Intelligence (BI) Analysis:

1. **Data Exploration and Visualization:**
   - Create interactive dashboards using BI tools to visualize trends and patterns in employee turnover.
   - Identify departments, roles, and specific projects with the highest turnover rates.
2. **Descriptive Analytics:**
   - Generate reports that highlight the primary reasons for attrition based on employee feedback, exit interviews, and other relevant sources.
   - Analyze the impact of factors like job satisfaction, workload, and career growth on employee turnover.
3. **Predictive Analytics with BI:**
   - Build predictive models within the BI tools to estimate the likelihood of turnover for current employees.
   - Implement scenario analysis to understand the potential impact of changes in satisfaction levels, compensation, or management practices.

Machine Learning Model:

1. **Data Preprocessing:**
   - Incorporate real-time data feeds from HR systems to ensure the model is continuously updated.
   - Dynamically handle new employee entries and update the model as employees leave or join.
2. **Feature Engineering:**
   - Include features such as recent performance reviews, project completion milestones, and employee engagement scores for a more accurate prediction.
3. **Model Training and Monitoring:**
   - Implement a mechanism to retrain the machine learning model periodically with the latest data.
   - Set up monitoring to alert HR teams when an employee's predicted turnover likelihood surpasses a certain threshold.

4. **Integration with BI Tools:**
    - Embed live predictions from the machine learning model into the BI dashboards.
    - Enable HR managers to drill down into specific departments or teams to identify high-risk individuals and take proactive measures.

Real-time Scenarios and Impact:

1. **Proactive Employee Retention:**
    - HR managers can use the integrated BI tools to identify high-risk employees and take proactive measures to address their concerns.
    - Real-time alerts enable timely interventions, such as personalized career development plans or targeted retention efforts.
2. **Strategic Workforce Planning:**
    - HR leaders can leverage predictive analytics to inform strategic workforce planning, ensuring that teams critical to ongoing projects are adequately supported.
3. **Continuous Improvement:**
    - Regular updates to the machine learning model based on real-time data allow for continuous improvement in prediction accuracy.
    - Feedback loops from HR teams can be integrated into the model to enhance its effectiveness over time.

By addressing the live problem statement of employee turnover at Acme Corporation, this project integrates BI tools and machine learning to provide actionable insights and empower the organization to proactively manage its workforce. The real-time nature of the analysis ensures that decision-makers have up-to-date information for effective interventions.

Next
Dashboards Buiding

Last updated 3 months ago