# **BUSINESS CASE**

## **INX Future Inc Employee Performance**

INX Future Inc, (referred as INX), is one of the leading data analytics and automation solutions provider with over 15 years of global business presence. INX is consistently rated as top 20 best employers past 5 years. INX human resource policies are considered as employee friendly and widely perceived as best practices in the industry.

Recent years, the employee performance indexes are not healthy and this is becoming a growing concern among the top management. There has been increased escalations on service delivery and client satisfaction levels came down by 8 percentage points.

CEO, Mr. Brain, knows the issues but concerned to take any actions in penalizing non-performing employees as this would affect the employee morale of all the employees in general and may further reduce the performance. Also, the market perception best employer and thereby attracting best talents to join the company.

Mr. Brain decided to initiate a data science project, which analyses the current employee data and find the core underlying causes of this performance issues. Mr. Brain, being a data scientist himself, expects the findings of this project will help him to take right course of actions. He also expects the clear indicators of non-performing employees, so that any penalization of non-performing employee, if required, may not significantly affect other employee morals.

# **PROJECT OVERVIEW:**

The goal of this project is to analyse employee performance within the organization and develop predictive models to enhance hiring processes and improve overall performance.

**The project will focus on four key objectives:**

1. Department-wise Performances Analysis:

2. Identification of Top 3 Important Factors Affecting Employee Performance:

3. A trained model which can predict the employee performance based on factors as inputs. This will be used to hire employees

4. Recommendations for Performance Improvement

## **Problem Statement:**

* The organization faces challenges in accurately assessing employee performance, leading to inefficiencies in hiring processes and suboptimal workforce productivity.
* To address these challenges, the aim is to develop a data science project which identifies the root causes of declining employee performance indices at INX Future Inc. and provide clear indicators for non-performing employees without negatively impacting overall employee morale or the company's reputation as a top employer.

# **DOMAIN ANALYSIS:**

**EmpNumber:** Represents a unique identifier for each employee. It serves as a primary key for identifying individual employees in the dataset.

**Age:** Represents the age of each employee. It is essential for analyzing demographics, workforce distribution, and understanding age-related factors affecting performance or attrition.

**Gender:** Indicates the gender of each employee. Gender diversity analysis can provide insights into gender balance within the organization and its impact on various aspects such as job roles, promotions, and salaries.

**EducationBackground:** Describes the educational background of employees.

**MaritalStatus:** Indicates the marital status of employees.

**EmpDepartment:** Specifies the department in which each employee works. Analysis of department-wise performance, attrition rates, and other metrics can provide insights into the organization's structure and functioning.

**EmpJobRole:** Describes the specific job role or position held by each employee within their department.

**BusinessTravelFrequency:** Indicates how frequently employees travel for business purposes. Analysis of travel patterns can help in managing travel-related expenses, work-life balance considerations, and understanding job demands.

**DistanceFromHome:** Represents the distance between an employee's home and the workplace. It can be relevant for assessing commuting challenges, relocation opportunities, and their impact on employee satisfaction and retention.

**EmpEducationLevel:** Specifies the education level attained by each employee, ranging from 'Below College' to 'Doctor'.

**EmpEnvironmentSatisfaction:** Indicates the level of satisfaction employees have with their work environment, ranging from '1' to '4'.

**EmpHourlyRate:** Represents the hourly rate of pay for each employee.

**EmpJobInvolvement:** Describes the level of involvement or engagement employees have in their jobs, ranging from '1' to '4'.

**EmpJobLevel:** Specifies the hierarchical level or rank of each employee's job within the organization.

**EmpJobSatisfaction:** Represents the level of satisfaction employees have with their jobs, ranging from '1' to '4'.

**NumCompaniesWorked:** Represents the number of companies employees have worked for previously.

**OverTime:** Indicates whether employees work overtime or not.

**EmpLastSalaryHikePercent:** Specifies the percentage of the last salary hike received by each employee.

**EmpRelationshipSatisfaction:** Indicates the level of satisfaction employees have with their work relationships, ranging from '1' to '4'.

**TotalWorkExperienceInYears:** Represents the total work experience (in years) of each employee.

**TrainingTimesLastYear:** Specifies the number of training sessions attended by employees in the last year. Analysis of training effectiveness and professional development opportunities can be performed using this column.

**EmpWorkLifeBalance:** Indicates the perceived level of work-life balance by employees, ranging from '1' to '4'.

**ExperienceYearsAtThisCompany:** Represents the number of years employees have been working at the current company.

**ExperienceYearsInCurrentRole:** Specifies the number of years employees have been in their current job role.

**YearsSinceLastPromotion:** Indicates the number of years since the last promotion received by employees.

**YearsWithCurrManager:** Represents the number of years employees have been working with their current manager.

**Attrition:** Indicates whether employees have left the company (attrition) or are still employed.

**PerformanceRating:** Specifies the performance rating assigned to each employee, ranging from '2' to '4'.

# **ANALYSIS METHODOLOGY**

Both visual analysis and statistical methods are carried out for data exploration.

## **Visual analysis methods:**

* Univariate analysis
* Bivariate analysis
* Multivariate analysis

## **Statistical methods used:**

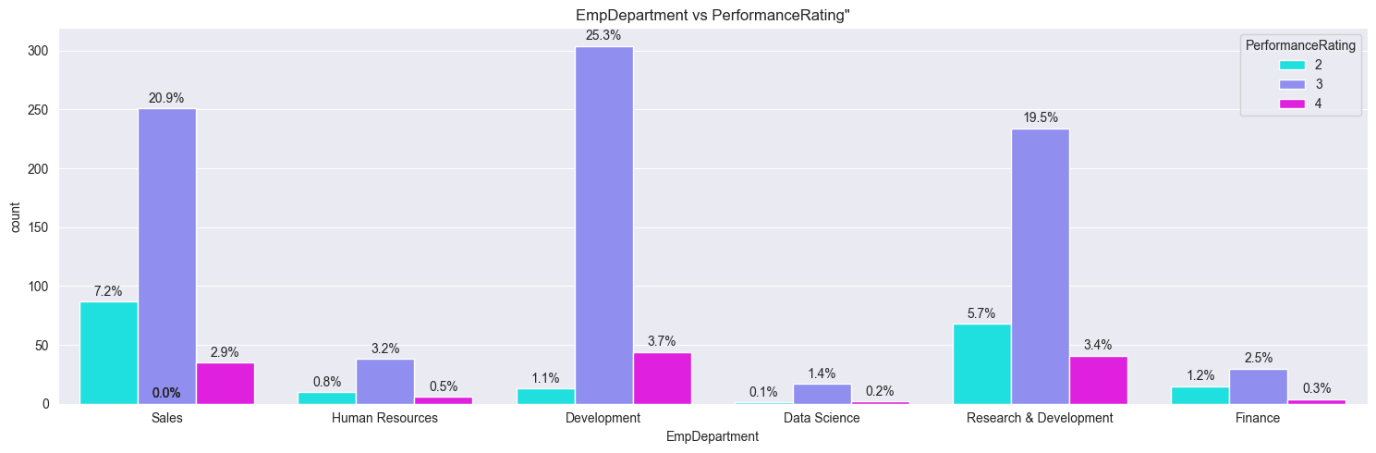
* Correlation coefficient

## **Insights driven from the analysis:**

* **Gender**: The distribution of performance ratings across genders can reveal if there is a gender bias in performance evaluations.
* **EducationBackground**: Understanding how employees from different educational backgrounds are rated can help in assessing if educational diversity impacts performance perceptions.
* **MaritalStatus**: This plot can show if marital status influences stability and performance ratings, potentially indicating if personal life stability translates into professional performance.
* **EmpDepartment**: Different departments might have varying benchmarks for performance ratings. This plot helps in identifying if certain departments rate their employees more stringently or leniently.
* **EmpJobRole**: Similar to departments, this can show if specific job roles are associated with higher or lower performance ratings, which might reflect on the expectations and pressures associated with those roles.
* **BusinessTravelFrequency**: Employees who travel more or less frequently might experience different stress levels and work challenges, which could affect their performance ratings.
* **EmpEducationLevel**, **EmpEnvironmentSatisfaction**, **EmpJobInvolvement**, **EmpJobLevel**, **EmpJobSatisfaction**: These plots can highlight how intrinsic job factors and personal employee satisfaction levels correlate with performance ratings.
* **NumCompaniesWorked**: Insights from this plot can indicate if having experience in multiple companies affects performance positively or negatively.
* **OverTime**: Overworking can either be seen as a sign of dedication or a route to burnout. This plot can help understand how overtime is affecting employee performance ratings.
* **EmpLastSalaryHikePercent**: This could show if salary hikes are aligned with performance ratings, potentially indicating if financial rewards are being used effectively as a motivational tool.
* **EmpRelationshipSatisfaction**, **TrainingTimesLastYear**, **EmpWorkLifeBalance**: These factors contribute to an employee's overall work satisfaction and could directly impact their performance ratings.
* **ExperienceYearsInCurrentRole**, **YearsSinceLastPromotion**, **YearsWithCurrManager**: These plots can provide insights into career progression and its impact on performance ratings.

## **Department wise performance analysis:**

The analysis assesses the performance of employees across different departments to identify any disparities or trends.



**Insights:**

* Development department dominates the other departments in number.
* In terms of performance rating, employees who belong to development department perform well with the dominating 25% of rating 2.
* Hence departments such as **development, sales and Research & development** constitute the high performing employees.

## **Top 3 factors affecting the employee’s performance:**

From the data analysis and the feature engineering, there are some of the factors which proved crucial in predicting the employee’s performance.

**Among them the top three factors are as factors:**

1. EmpEnvironmentSatisfaction
2. EmpLastSalaryHikePercent
3. EmpDepartment\_Development

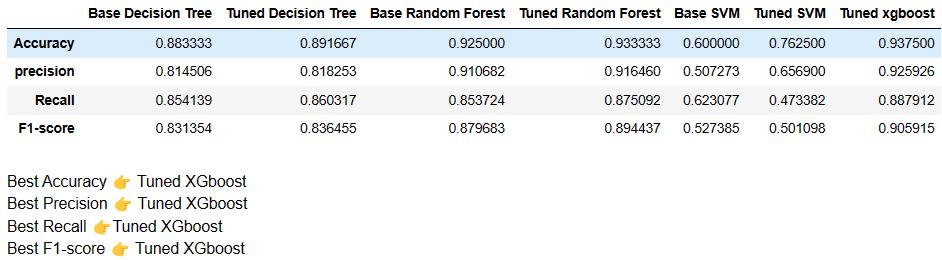
# **Algorithms used in this project:**

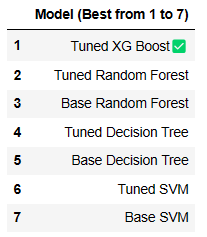
* Decision Tree classifier
* Random Forest classifier
* Support Vector Machine
* Extreme Gradient Boosting

## **Trained models:**

1. **DECISION TREE CLASSIFIER**
2. **TUNED DECISION TREE CLASSIFIER**
3. **RANDOM FOREST CLASSIFIER**
4. **TUNED RANDOM FOREST CLASSIFIER**
5. **SUPPORT VECTOR MACHINE**
6. **TUNED SUPPORT VECTOR CLASSFIER**
7. **TNED XG BOOSTING**

**Model results:**





# **RECOMMENDATION:**

## **Model recommendation:**

* The top three would be
  + Tuned XG Boost model
  + Tuned Random forest model
  + Base Random Forest model
* These three models achieved high accuracy scores of range 93 - 90% indicating strong predictive power. However, the choice of the model may depend on other factors, including resource constraints and model interpretability.
* **XG Boost:** Strongly recommended, it’s efficient and highly accurate, but may require fine-tuning.
* **Random Forest:** Recommended if computational resources are available.

## **Recommendations for business problem:**

1. Prompt incentives and acknowledgement: Since employees' performance is directly impacted by their salary boost %, it is imperative to track their progress and make sure that their effort receives the right recognition and a fair raise.   
  
2. Open communication: Since it has been demonstrated that an employee's performance ratings are greatly influenced by their environment, promptly and empathetically addressing employee complaints can help increase output and productivity.

3. Establishing a friendly and inclusive workplace: Cooperation and teamwork should promote improved departmental performance in all job roles.   
  
4. Regular performance reviews: to educate employees of their progress and motivate them to make changes in challenging areas.

5. Training initiatives: Programmes for the timely and pertinent development of skills

## **Conclusion:**

To improve employee performance, reasonable and caring initiatives are required. The aforementioned recommendations must be followed in a methodical and planned manner to guarantee that the right results are expected in a fair amount of time. The top-down hierarchy may also have an effect on the morale of junior employees; therefore, HR collaborations with staff members will be helpful in identifying a decentralised but effective solution to this issue.