Improving Employee Retention by Predicting Employee Attrition Using Machine Learning

Rakamin
Academy

Created by:
Dinda Galuh
dindagaluhg@gmail.com
linkedin.com/in/dinda-galuh-guminta

Statistics graduate with passion in data visualization, data analysis, and reporting. Dedicated and hard working person.

I joined the data science bootcamp because I am more interested and have a passion for learning and working in the data field. I like doing data analysis and data visualization. I want to apply my knowledge in the real world.

Human resources (HR) are the main assets that need to be managed well by the company so that business goals can be achieved effectively and efficiently. On this occasion, we will face a problem regarding human resources in the company. Our focus is to find out how to keep employees staying in the current company which can result in increased costs for employee recruitment and training for those who have just joined. By knowing the main factors that cause employees to feel uncomfortable, companies can immediately overcome them by creating programs that are relevant to employee problems.

Data Preprocessing

- **Handle missing value**
  - Fill missing value in AlasanResign column with "masih_bekerja"
  - Fill missing value in JumlahKetidakhadiran, JumlahKeikutsertaanProjek, and JumlahKeterlambatanSebulanTerakhir column with 0 (zero)
  - Drop row with missing value in SkorKepuasanPegawai column
- **Check value in categorical column**
  - Replace value "Product Design (UI & UX)" with "masih_bekerja" in AlasanResign column
  - Replace value "1" with "yes" in PernahBekerja column
  - Replace value "–" with "NaN" in TanggalResign column
- **Change data type that contain date**
- **Drop unused columns**
  - Drop PernahBekerja column because it contains 1 value (constant)
  - Drop IkutProgramLOP column because it has many missing value (about 89.9% missing)

Untuk selengkapnya, dapat melihat jupyter notebook disini

Annual Report on Employee Number Changes
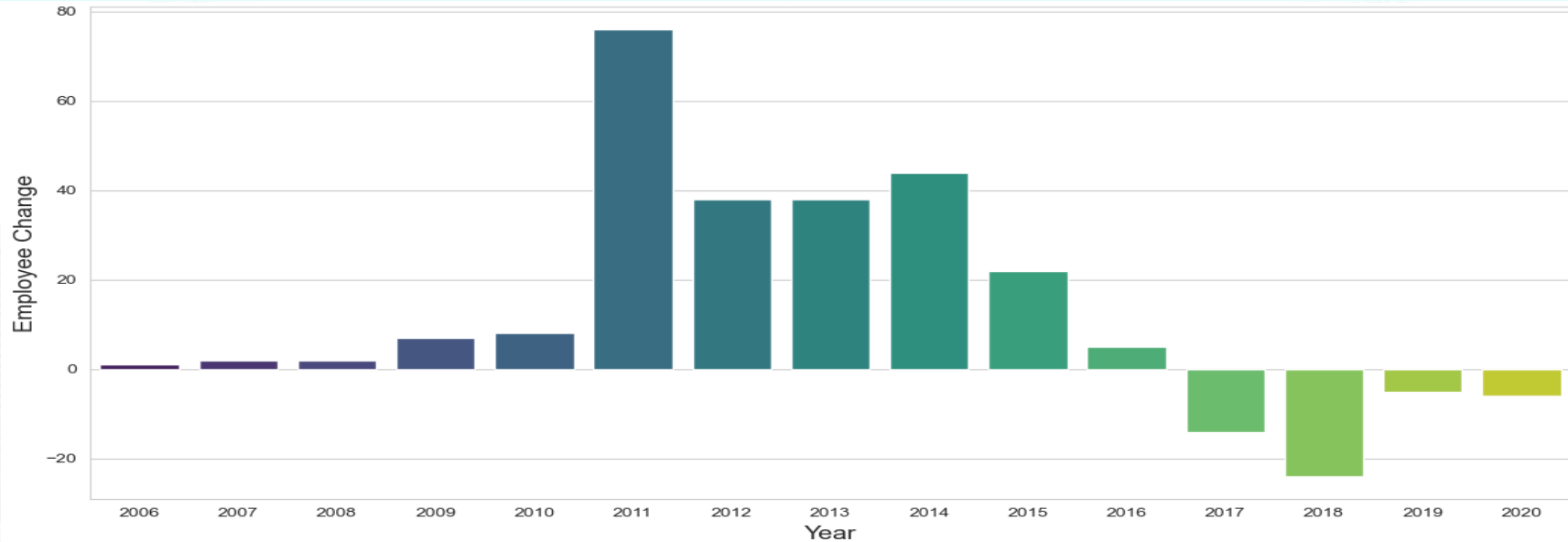
Early Growth (2006-2010):
It appears that the company experienced steady growth in employee recruitment, with the addition of 1-8 employees each year.

Hiring Surge (2011-2014):
- A huge spike in new employee hiring. This could indicate a major expansion or major project the company is taking on.
- Although several employees left in 2013 and 2014, the company still experienced net growth in the number of employees in this period.

Hiring Slowdown (2015-2016):
Employee growth slows indicate a phase where the company may focus on training, internal development, or integration of previously recruited new employees.

Employee Losses & Potential Issues (2017-2020):
- The company experienced a decline in the number of employees every year.
- A drastic decline, in 2017 - 2018, is a sign that there is a problem at the company, perhaps related to work culture, management, or unfavorable market conditions.
- The absence of new hires in 2019 - 2020 also indicates that the company is not in a position to grow or it is possible that they are choosing not to hire to save costs.

Percentage of Employees Not Resigned by Job Role
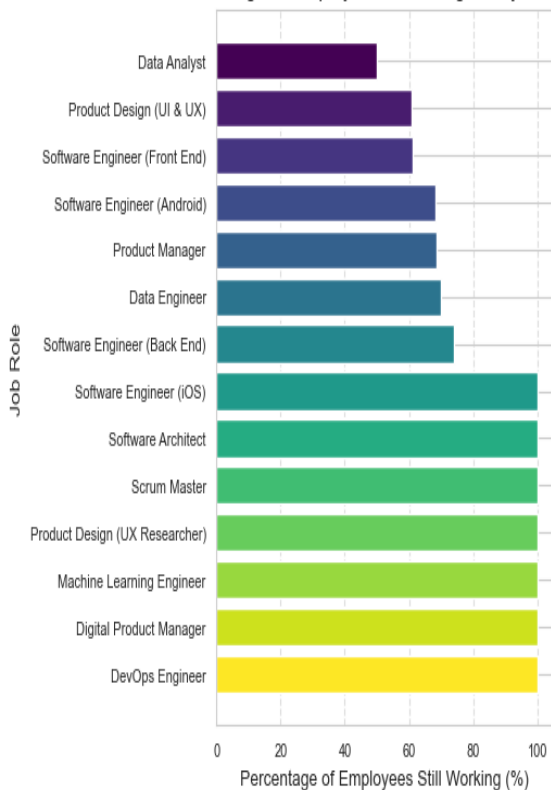
Percentage of Employee Resignation by Reason

Toxic Culture as the Main Reason: a strong indication that the data analyst work culture does not support employee welfare.

Resign Reason Analysis for Data Analyst Job Role

- **Achievement and Reasons for Resigning**: This shows that a toxic work culture affects employees at all levels of achievement.
- **Fresh Graduates are More Affected**: Employees who have just entered the workforce are more susceptible to the effects of a toxic work culture, or less tolerant of such a culture compared to more experienced employees.

Recommendation:
- **Work Culture Review**: Management must immediately conduct an in-depth study of what causes work culture to be considered "toxic". Using employee surveys, interviews, and focus groups can help identify key problems.
- **Orientation and Coaching Program**: Fresh graduates need additional guidance and support when first entering the workforce. Providing an effective orientation or mentoring program can help them adapt to the work environment.
- **Improve Communication**: Sometimes, the perception of a "toxic culture" can arise from misunderstandings or lack of communication. Make sure that all employees feel heard and that they understand the company's goals and values.
- **Review HR Processes**: Ensure that the HR department has a proper process for handling employee complaints.

The features used in the classification model are:

- Status Kepegawaian
- Performance Pegawai
- Pekerjaan
- Skor Kepuasan Pegawai
- Umur

In this case, the metric used is:

- Recall (Sensitivity): Indicates what percentage of employees who actually resigned was identified by the model. Recall is important to ensure that we don't miss many positive cases (employees who resigned).
- F1 Score: F1 score is the harmonic average between precision and recall. It's a good metric to balance between precision and recall. The F1 score is good if we want to strike a balance between identifying employees who are leaving and minimizing misidentification.
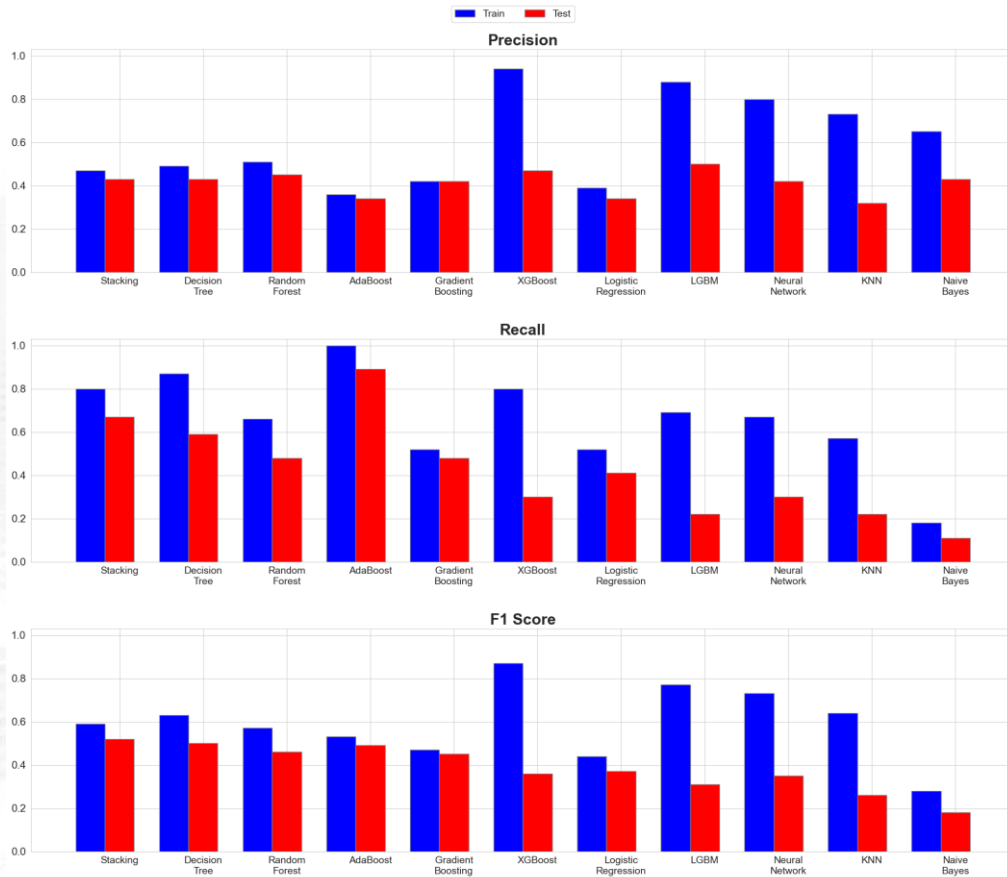
Precision Comparison

Recall Comparison

F1 score Comparison
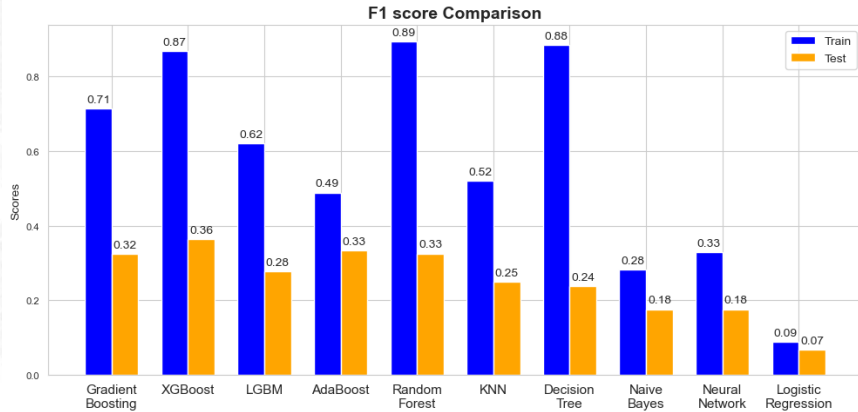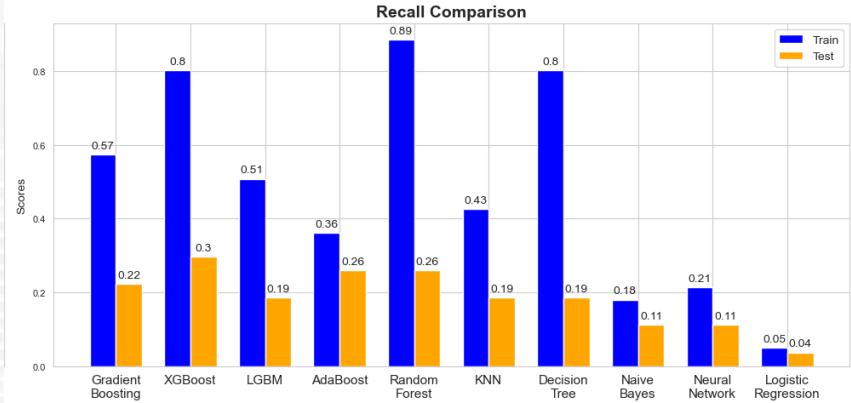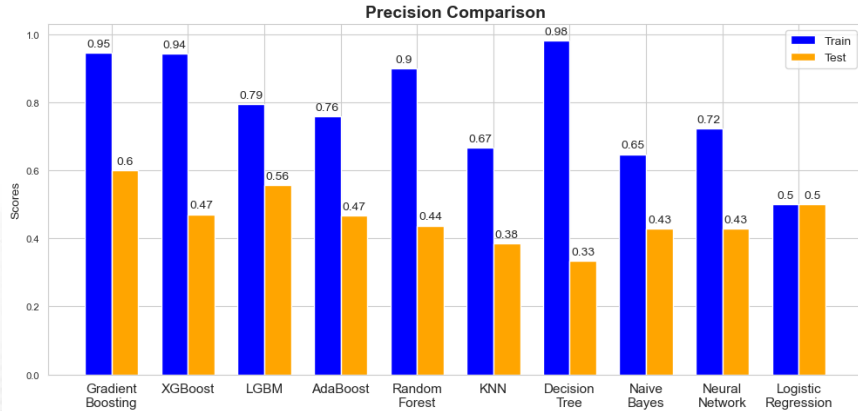
- A balanced model in terms of precision and recall on the test set, Gradient Boosting seems decent
- On the test set, Gradient Boosting, XGBoost, and AdaBoost have the highest F1 scores.
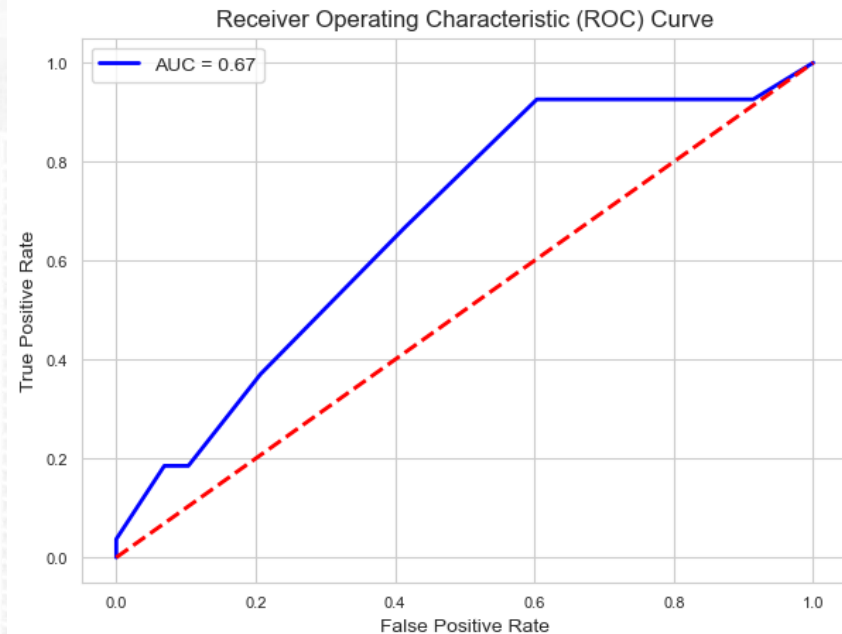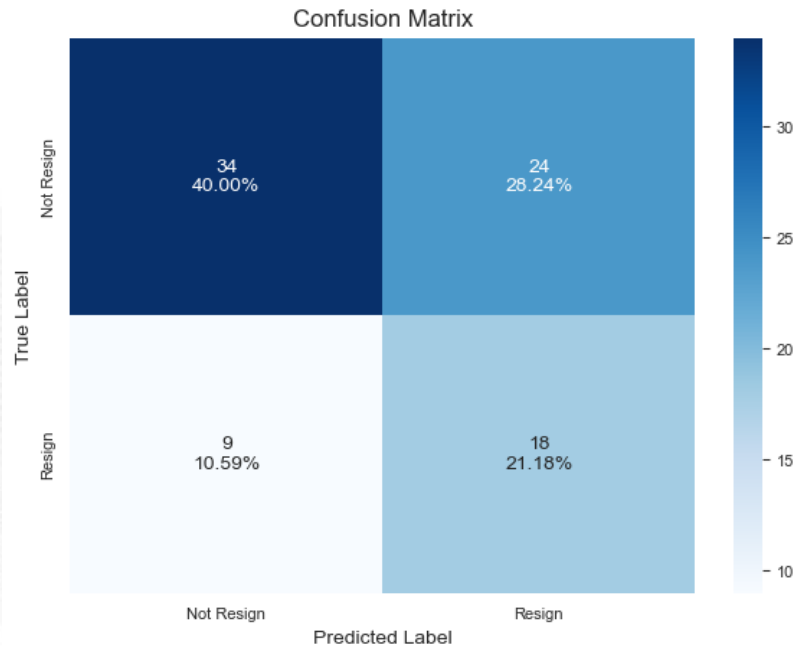
Model Performance after Hyperparameter Tuning

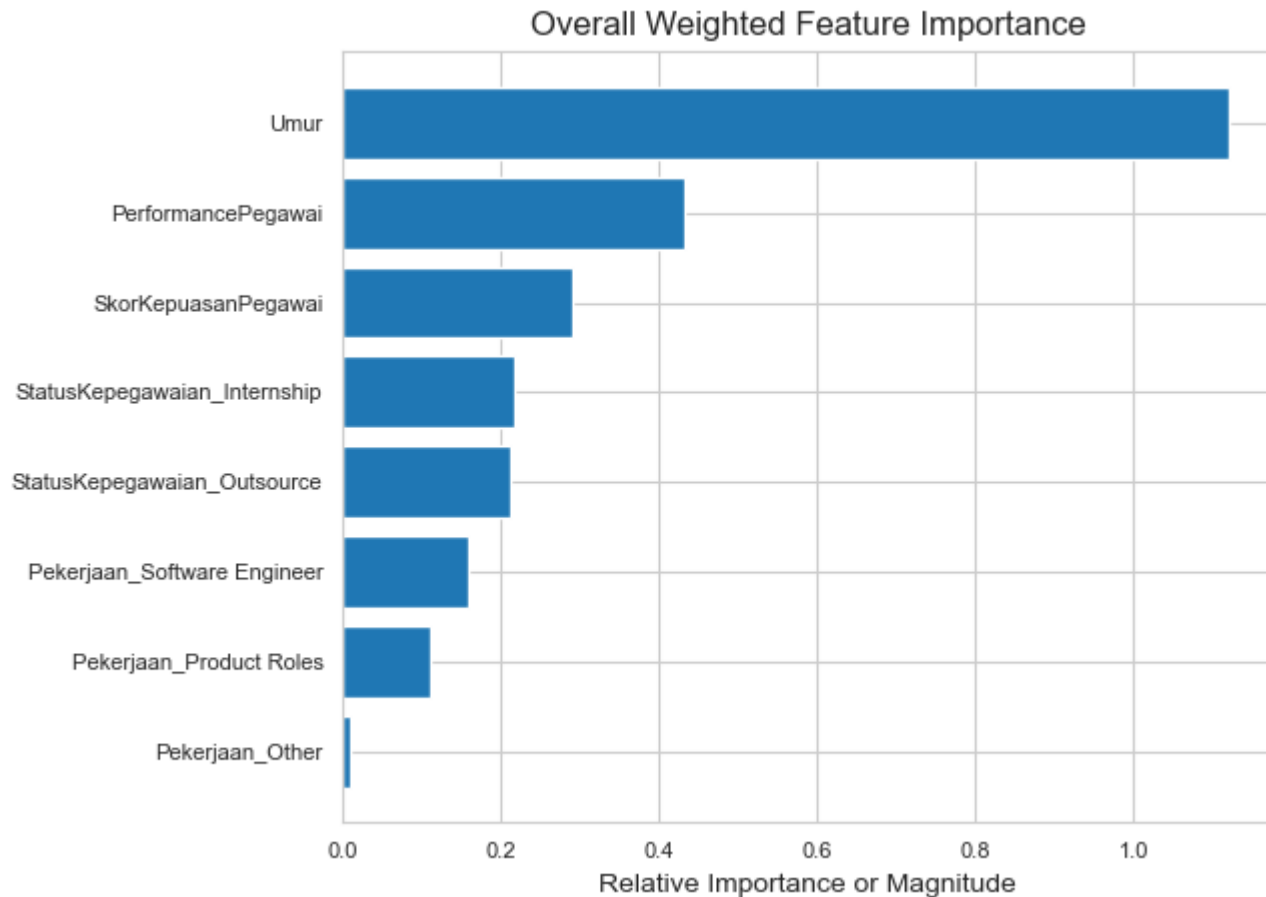The Stacking model shows solid results on recall and F1 scores on the training and test data

Model Performance using Best Method – Stacking

Confusion Matrix

	Not Resign	Resign
Not Resign	34 / 40.00%	24 / 28.24%
Resign	9 / 10.59%	18 / 21.18%

Predicted Label / True Label

Receiver Operating Characteristic (ROC) Curve

AUC = 0.67

Stacking – Feature Importance
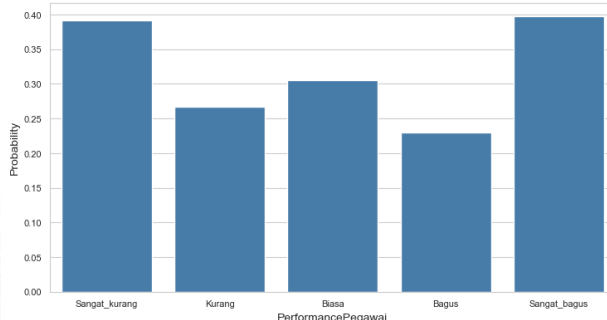

Overall Weighted Feature Importance

1. Age was a consistently important feature in the model.
2. Features such as Employee Performance and Employee Satisfaction Score also contribute greatly to the model.
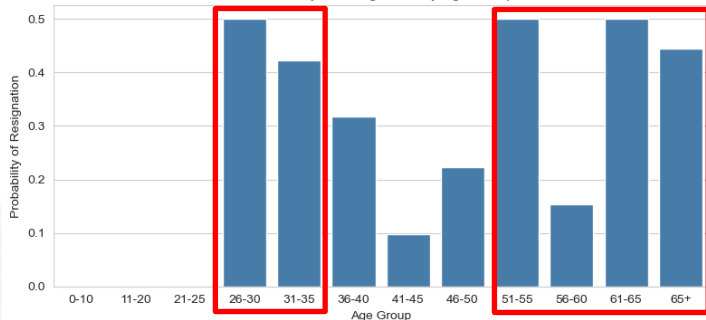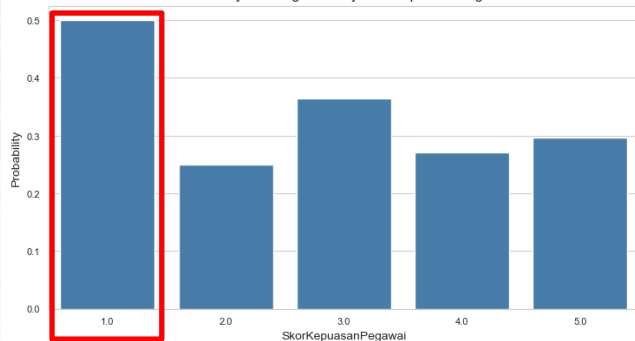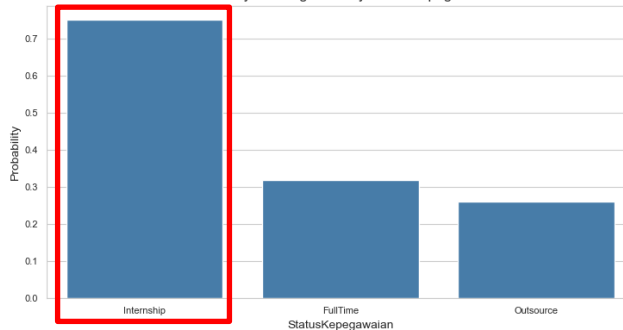
- Employees aged 26–35 and 51 years above have a higher chance of resigning
- Employees with a satisfaction score of 1 have the highest chance of resigning
- Employees with internship status have a chance of resigning up to 70%

Conclusion

In the modern era, where talent has become a valuable asset for companies, understanding the reasons behind an employee's decision to resign is very important. The results of this model offer insights that can help companies better understand what might drive their employees to leave the company.

- Age emerged as a consistently important feature in models. This could indicate that there are certain phases in a person's career where they are more likely to consider other options, perhaps when reaching a certain point in their career or as they approach retirement.

- Employee performance is also a significant factor. Employees with certain performance, whether excellent or unsatisfactory, may have a higher tendency to resign. High performers may be looking for better opportunities while low performers may feel like they don't fit in with the company.

- Employment Status Internship, emerged as highly significant features.

Recommendation

- **Considering Age**: Companies must understand the needs of employees at different phases of their careers. Mentorship programs and competitive compensation for younger employees and continued development opportunities for senior employees may be helpful.
- **Performance Management**: For employees who perform well, companies must ensure that they feel valued and have a clear career path. Meanwhile, for employees with unsatisfactory performance, training and coaching programs may be able to help.
- **Job Evaluation**: Further analyze the types of jobs or employment status identified by the model as risk factors. There may be a need to review compensation, workload, or working conditions for certain positions.
- **Employee Satisfaction Surveys**: Considering that employee satisfaction scores are an important factor, companies should regularly conduct employee satisfaction surveys to understand their needs and concerns.
- **Enhanced Onboarding Process**: The onboarding process for interns might not be as rigorous or informative as that for permanent employees. Enhancing this process can provide interns with a clearer understanding of their roles and expectations.

By implementing these recommendations, companies can minimize employee turnover and ensure that their employees feel satisfied and appreciated, which will ultimately contribute to the company's long-term success.