Institute of Data

Capstone Project Documentation

Employee Attrition

Samantha Carlson

May 2021

# Contents

[Contents 2](#_Toc73177084)

[1. Background 4](#_Toc73177085)

[1.1. Employee attrition 4](#_Toc73177086)

[1.2. Current state – associated costs 4](#_Toc73177087)

[1.3. Exceptions 4](#_Toc73177088)

[1.4. Desired state 1](#_Toc73177089)

[1.5. Other research projects 1](#_Toc73177090)

[2. Industry 1](#_Toc73177091)

[1.6. Current state 2](#_Toc73177092)

[1.7. Other industries. 2](#_Toc73177093)

[3. Stakeholders 2](#_Toc73177094)

[4. Business question 3](#_Toc73177095)

[1.8. Considerations 3](#_Toc73177096)

[1.9. Value 3](#_Toc73177097)

[1.10. Accuracy 2](#_Toc73177098)

[1.10.1. Project accuracy 2](#_Toc73177099)

[1.10.2. Modelling accuracy 2](#_Toc73177100)

[1.10.3. False positives/negatives 2](#_Toc73177101)

[5. Data question 3](#_Toc73177102)

[6. Data 3](#_Toc73177103)

[1.11. Overview 3](#_Toc73177104)

[1.12. Data reliability 3](#_Toc73177105)

[1.13. Data quality 2](#_Toc73177106)

[1.14. Ongoing access to data 2](#_Toc73177107)

[1.15. Data snapshot 2](#_Toc73177108)

[7. Data science process 2](#_Toc73177109)

[1.16. Data analysis 2](#_Toc73177110)

[1.16.1. Data pipeline 2](#_Toc73177111)

[1.16.2. Data cleaning 2](#_Toc73177112)

[1.16.3. . Exploratory Data Analysis – Target feature (Attrition) 2](#_Toc73177113)

[1.16.4. Addressing data imbalance 3](#_Toc73177114)

[1.16.5. Exploratory Data Analysis (Features vs. Attrition) 1](#_Toc73177115)

[1.17. Modelling 1](#_Toc73177116)

[1.17.1. Feature selection 1](#_Toc73177117)

[1.17.2. Feature correlations 1](#_Toc73177118)

[1.17.3. Subset of key features 1](#_Toc73177119)

[1.17.4. Model selection 1](#_Toc73177120)

[1.17.5. Training models 1](#_Toc73177121)

[1.17.6. Model performance (individual) 1](#_Toc73177122)

[1.17.7. Model performance (ensemble models) 1](#_Toc73177123)

[1.17.8. Model performance - ROC/AUC 1](#_Toc73177124)

[1.18. Implementation 2](#_Toc73177125)

[8. Data answer 1](#_Toc73177126)

[1.19. Answer 1](#_Toc73177127)

[1.20. Confidence 1](#_Toc73177128)

[9. Business answer 1](#_Toc73177129)

[1.21. Answer 1](#_Toc73177130)

[1.22. Confidence 1](#_Toc73177131)

[10. Response to stakeholders 1](#_Toc73177132)

[11. End-to-end solution 2](#_Toc73177133)

[12. Appendix 1 – Data snapshot 1](#_Toc73177134)

[13. Appendix 2 - Data dictionary 1](#_Toc73177135)

[14. Appendix 3 - References 1](#_Toc73177136)

# Background

## Employee attrition

Employee attrition refers to the number or percentage of workers who leave an organization and are replaced by new employees.

This includes any employee departure, including resignations, layoffs, terminations, retirements, location transfers, or even deaths.

## Current state – associated costs

Employee attrition can cost employers 33 percent of an employee’s annual salary. This cost comes down to several features but is primarily attributed to the cost of hiring of a replacement.

It is important to note that replacement costs are not simply recruitment fees, though these are significant and can cost anywhere between 20 and 30% of the annual salary of the position. The cost of replacing an employee also includes advertising the role, interview costs, as well as the cost of training the new hired employee, and arranging for the workload to be completed while the new employee learns the role.

There are several hidden, non-financial, costs associated with employee attrition. These unseen costs are from aspects required to ensure the organisation continues to operate and can include, the burden placed on others to pick up the additional work as well as the potential damage to morale and team dynamics.

It is important to note that not all employee attrition is bad. In fact, in some cases it is even beneficial, for example an underperforming employee. It can cost less to hire and train a new employee than to keep someone with poor performance. In addition to this, employee morale can increase with such a departure as the workload may become more evenly distributed if all employees are performing to the expected standard. Another positive reason for employee attrition is in the area of learning and development.; if an employee moves to another organisation for a promotion opportunity this can be considered healthy attrition. While this does still impose a cost on an organisation, it is expected and accepted that employees will move on if they are offered a more senior role elsewhere, this is particularly true in smaller organisations where the opportunities for promotion are not as abundant.

## Exceptions

This study will focus on the public, not-for-profit, service provision sector as outlined in [Section 2](#_Industry) so we will not discuss here the costs associated with a loss of revenue for a sales organisation.

## 

## Desired state

As outlined in the August 2018 AHRI report, participants surveyed confirmed that a desirable rate for attrition would be between 1% and 10%.

If we could predict the employees most likely to leave an organisation, we could minimise the associated costs – both financial and those that are unseen.

## Other research projects

There have been numerous other studies undertaken on this topic (both machine learning and other). There are too many individual studies to note here, but many of these are available online.

To support my research, those studies that I referenced the most are listed below, all other sources can be found in [Appendix 4 - References](#_Appendix_4_-):

1. [The Australian Human Resources Institute (AHRI) through the Turnover and Retention Research Report, August 2018](https://www.ahri.com.au/media/1222/turnover-and-retention-report_final.pdf)
2. [Victorian Public Sector Commission, Employee turnover and mobility.](https://vpsc.vic.gov.au/data-and-research/data-facts-visuals-state-of-the-sector/employee-turnover-and-mobility/)

This research tells us that employee attrition is an ongoing business issue to be addressed. While machine learning can help us to predict those employees who are more likely to stay/leave an organisation, the real work to address the problem is found in the subsequent actions taken, for example:

* What initiatives did the organisation implement to retain employees?
* Why do employees want to leave, and can this be addressed?

Finally, as there is a ‘healthy’ level of employee attrition, it is not only important to undertake these subsequent actions, but to follow up with further analysis and action implementation, ideally this is an ongoing, dynamic exercise that occurs regularly within organisations to meet employee needs.

# Industry

This study will focus on the Victorian Public Service (VPS), non-profit, service industry. For example, Court Services.

Public sector entities have a funding model whereby they are funded directly by the population that they serve, in other words – taxpayers. For these taxes paid, it is important that the government ensures a worthwhile return-on-investment for its citizens. There are several reasons for this, but I have highlighted two key ones below:

1. If an elected government is seen to mis-use public funding, or not provide value for money for its citizens, the likelihood of re-election is diminished.
2. If public funds are not seen by the public to be supporting worthwhile ventures, there will be public dissatisfaction with the elected government, which can lead to unrest, protests, complaints, etc.

## Current state

For the 2019-20 financial year, the VPS had an attrition rate of 13%.

As of June 2020, the VPS employment totalled 266,272 in full-time equivalent employees, this made up 9% of the Victorian labour force, and included over 1,820 employers.

If we consider that the median salary over 2019-20 for a VPS employee was $95,549 and calculate attrition costs at the rate of 13%, this is a potential cost to the taxpayer of over $3 billion.

## Other industries.

While this project focusses on the VPS and taxpayer funding, the concepts considered could be applied to other industries with some modification. For example, the inclusion of sales revenue and potential losses, as well as other factors that may influence an employee’s decision to leave their current position, for example, key performance indicators and targets.

# Stakeholders

The stakeholders for this project and their levels of interest are detailed below in [Table 1](#Table1).

**Table 1: Employee attrition stakeholders**

| **Stakeholder** | **Interest** | **Expectation** |
| --- | --- | --- |
| Taxpayers | That VPS operations run to budget and do not cost the taxpayer additional amounts through the increase of taxes.  That there is a value-for-money return on the investment into the VPS made by taxpayers. | That taxes do not increase.  That taxpayer funding is used efficiently.  The VPS services are delivered and provide a worthwhile return on investment. |
| Current government representatives | That policies and initiatives are delivered within budget to increase the likelihood of a re-election. | Projects are completed within budget.  Projects and VPS entities that require funding provide the service expected and deliver project outcomes.  Funded projects meet campaign promises and budget. |
| Victorian Public Service entities | That there is adequate budget to resource operations and initiatives.  That there are sufficient employees to undertake the work of the organisation. | That the state budget provides adequate funding for operations.  That there is a fair and considered approach to the allocation of public funding. |
| Department of Treasury and Finance | That all finances allocated are spent in an appropriate manner. | That operations do not exceed the allocated budget. |
| Private sector business owners | If the cost of public service operations increases there will likely be an increase in taxes, which will impact their own business profits and finances. | The VPS will monitor and be accountable for all public sector employee related expenditure.  The VPS Executives will monitor their budgets to ensure no overspend. |

# Business question

Can we predict which employees are most likely to leave an organisation?

## Considerations

To answer this business question, I will consider:

* Which factors that can be addressed lawfully by an employer most impact employee attrition?
  + Factors that are excluded from this project are those that are protected attributes as defined by the [Fair Work Act](https://www.fairwork.gov.au/employee-entitlements/protections-at-work/protection-from-discrimination-at-work). These include:
    - race
    - colour
    - sex
    - sexual orientation
    - age
    - physical or mental disability
    - marital status
    - family or carer's responsibilities
    - pregnancy
    - religion
    - political opinion
    - national extraction
    - social origin.

## Value

If a VPS employer can predict which employees are most likely to leave the organisation, they can better plan for the use of their financial resources. For example, does the employee who wishes to leave have reasons that can be resolved, or do they fit into the ‘good’ criteria for wishing to move on as discussed [above](#_Current_state_–)?

By reducing overall employee attrition to a more desirable rate of 10% or less, an organisation will reduce its spend on employee attrition activities (recruitment, etc.) and will be better placed to address the needs of their employees. For example, an uneven workload distribution.

For taxpayer and private business owner stakeholders, a reduced attrition rate of less than 10% would decrease the likelihood of rising taxes (for this reason, however this project does not address other reasons taxes may increase) and would ensure a more appropriate use of taxpayer funded activities. For example, an increase in the number of Court hearings compared to lower hearing rates due to the funding requirements of employee attrition.

For the current government and the Department of Treasury and Finance, this could increase the likelihood of a re-election and would decrease the potential reputational damage of an overspent budget.

## Accuracy

### Project accuracy

For this project, there is no specific, one-point-in-time accuracy rate. This project is to determine which employees are most likely to leave an organisation, and how this could be addressed to:

1. Enable an organisation to reduce the cost of employee attrition, and
2. Enable an organisation to reduce its employee attrition rate to under 10% as discussed [above](#_Desired_state).

Ideally, the model that has been created would be used on a regularly scheduled basis (for example, following the annual engagement survey or following the completion of employee professional development plan reviews).

It is important to note that this project accuracy relies heavily on the honesty and openness of employees, which in turn determines the level of accuracy of the data itself.

### Modelling accuracy

Given that we are trying to predict human behaviour and that we have a limited number of features that describe an employee’s situation, it seems unrealistic to expect a 90-100% accuracy for this type of modelling.

Based on the factors available that we can lawfully use to address employee performance and the likelihood of them staying/leaving an organisation the accepted accuracy level for this machine learning would be between 80-90%.

### False positives/negatives

Recommendations from this project (refer to [Response to stakeholders](#_Response_to_stakeholders)) suggest ongoing operational process improvements to improve employee retention rates. From these recommendations and the machine learning, false positive predictions (a prediction where an employee is incorrectly classified as likely to leave the organisation) will have no negative impacts on either the employee, or organisation.

While false negative rates (an employee is incorrectly classified as likely to stay with the organisation) are not ideal, if operational initiatives are put in place to improve overall employee retention, these employees will still be positively impacted by any changes made and will hopefully choose to remain with the organisation.

# Data question

Which factors most contribute to employee attrition?

To answer this question, we will require data that includes details of 1) Employee performance, and 2) Employee engagement.

# Data

## Overview

The data set used for this project was sourced from [data.world](https://data.world/aaizemberg/hr-employee-attrition), was generated by IBM, contains 1470 rows and 21 columns of information, divided as detailed below.

* Age
* Attrition
* Business Travel
* Daily Rate
* Department
* Distance from Home
* Education
* Education Field
* Employee Count
* Environment Satisfaction
* Gender
* Hourly Rate
* Job Involvement
* Job Level
* Job Role
* Job Satisfaction
* Marital Status
* Monthly Income
* Monthly Rate
* Number Companies Worked
* Over18
* Over Time
* Percent Salary Hike
* Performance Rating
* Relationship Satisfaction
* Standard Hours
* Stock Option Level
* Total Working Years
* Training Times Last Year
* Work Life Balance
* Years at Company
* Years in Current Role
* Years Since Last Promotion
* Years with Current Manager

## Data reliability

As human resources (employee) data is quite hard to come by due to Privacy Laws, IBM created this fictional data set to enable students to practice machine modelling with data typically found in organisations.

As this is a fictional dataset it is difficult to ascertain its reliability. However, this project is designed to be used by any organisation and will be adopted to suit the data that is available in each use case.

## Data quality

The raw data is of a high quality with most attributes already in the preferred integer format. There are no missing values.

## Ongoing access to data

For this project, the data used was a snapshot of information from a specific time/place. If this modelling were to be adopted by an organisation the data would be regularly available to enable ongoing reviews of employee statistics and attrition rates.

## Data snapshot

For a snapshot of the complete dataset, refer to [Appendix 1](#_Appendix_1_–).

For details of the input data, refer to [Appendix 2](#_Appendix_2_-).

# Data science process

## Data analysis

### Data pipeline

The data pipeline used is detailed below in [Figure 1](#Figure1).

**Figure 1: Data pipeline**

If this process and modelling were to be adopted for future use in organisations, this same data pipeline could be applied with modifications as required.

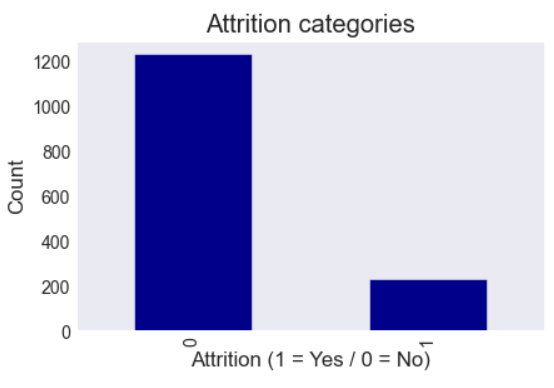
### Data cleaning

As this data set contained attributes that are protected by the Fair Work Act as outlined in 1.7 above, I have removed these features. In addition to this, there was a requirement to change the data types of some features from objects to integers to enable the categorical modelling to produce results. This dataset was otherwise clean.

### Exploratory Data Analysis – Target feature (Attrition)

[Figure 2](#Figure2) displays the distribution of data for the target feature, attrition.

**Figure 2: Distribution of target data**



### Addressing data imbalance

**General**

As we can see, the target feature columns are unbalanced. To address this, I have used the ‘stratify’ function when splitting my data into the train and test sets. In addition to this, I have used kFold validation to confirm my results, with 5 x splits.

**Specific - Stacking**

To address the imbalance within each stacking algorithm, I have weighted the classes for each model to ensure that they are given equal consideration with each modelling method.

**Specific-Boosting and Bagging**

Due to the nature of bagging and boosting techniques with decision trees, I have not weighted the classes for these two algorithms

### Exploratory Data Analysis (Features vs. Attrition)

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

As we can see from these visualisations, it is quite hard to view (at-a-glance) which features most impact attrition rates. For more descriptive visualisations, refer to section [Modelling](#_Modelling) to see feature correlations.

## Modelling

### Feature selection

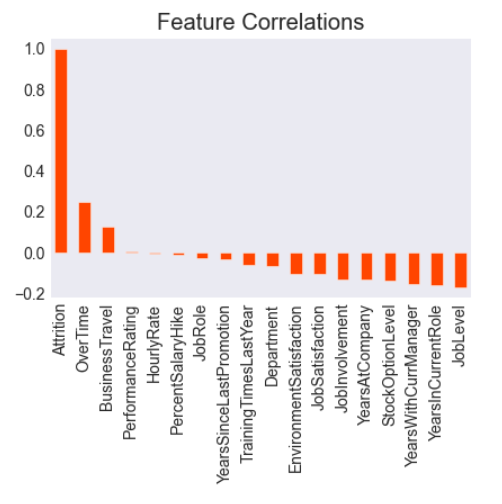
The final features selected to commence modelling were:

* Attrition
* Business Travel
* Department
* Environment Satisfaction
* Hourly Rate
* Job Involvement
* Job Level
* Job Role
* Job Satisfaction
* Over Time
* Percent Salary Hike
* Performance Rating
* Stock Option Level
* Training Times Last Year
* Years at Company
* Years in Current Role
* Years Since Last Promotion
* Years with Current Manager

### Feature correlations

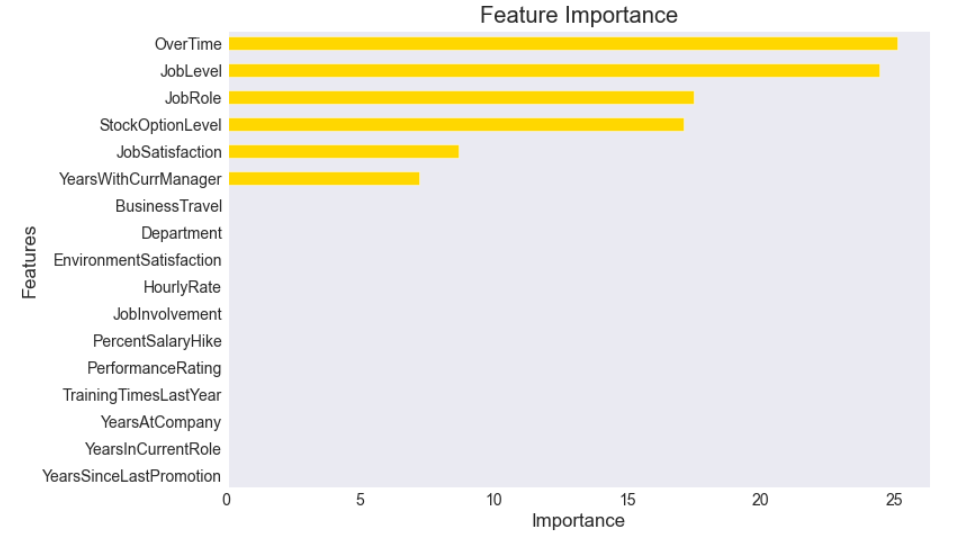
Refer to [Figure 3](#Figure3) below to view the features vs. target correlations.

**Figure 3: Feature correlations**



### Subset of key features

**Figure 4: Feature importance**



From [Figure 4](#Figure4) above, we can see that the features with the highest levels of importance are: Over time, job level, job role, stock option level, job satisfaction, and years with current manager.

After removing these features and rerunning my models, the performance dropped. From this we can assume that while the remaining features do not contribute as much to the model as the top six features, they do still play a part in determining whether an employee will leave the organisation. For this reason, I will be using all features to run my models.

### Model selection

As all features of this dataset are categorical, the models selected need to be able to process this. As such, the individual models chosen are:

* CatBoost
* XGBoost
* K-Nearest Neighbour
* Random Forest
* Light GBM
* Logistic Regression.

These models have then been combined into an ‘ensemble’ to produce my final results (refer to section [Model performance (ensemble models)](#_Model_performance_(ensemble)).

### Training models

As this is a relatively small dataset:

* It takes less than one minute to train the models, and
* All data and notebooks can be saved locally on the computer (though for backup reasons I also have copies saved on a cloud platform).

### Model performance (individual)

As discussed above in [False positives/negatives](#_False_positives/negatives), there are no significant consequences should an employee be incorrectly classified as likely to leave/remain with an organisation. As such, I will be using the accuracy measurement to test the reliability of my models instead of precision, recall, and F1. [Table 2](#Table2) below displays the accuracy scores for each of the individual models used.

**Table 2: Individual models and accuracy scores**

|  |  |
| --- | --- |
| **Model** | **Accuracy** |
| CatBoost | 0.76 |
| XGBoost | 0.85 |
| K-Nearest Neighbour | 0.83 |
| Random Forest | 0.86 |
| Light GBM | 0.84 |
| Logistic Regression | 0.79 |

### Model performance (ensemble models)

Following the individual modelling demonstrated above, the ensemble modelling was compiled. Results from each of the three options (Bagging, Boosting, Stacking) are detailed below in [Table 3](#Table3).

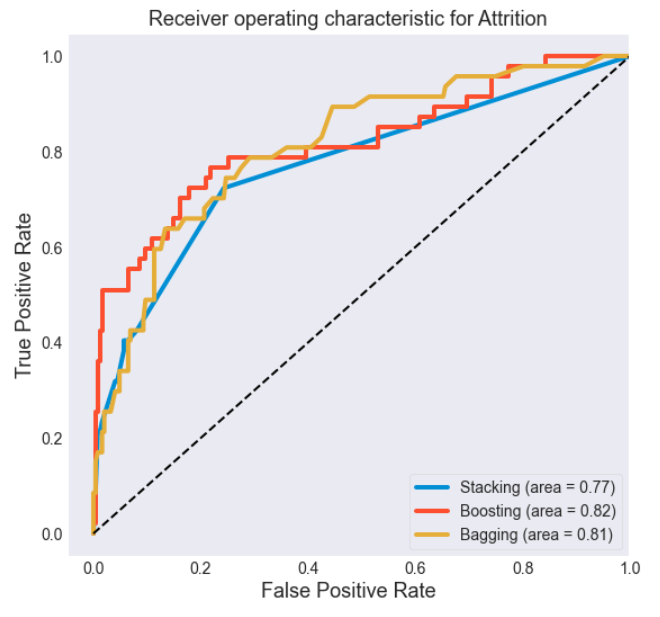
**Table 3: Ensemble model results**

|  |  |
| --- | --- |
| **Ensemble Model** | **Accuracy** |
| Stacking | 0.85 |
| Boosting | 0.89 |
| Bagging | 0.85 |

### Model performance - ROC/AUC

[Figure 5](#Figure5), below, demonstrates the ROC/AUC curve characteristics for each of the model results discussed above in [Table 3](#Table3).

**Figure 5: ROC/AUC curve (Bagging, Boosting, Stacking)**



## Implementation

To implement this model within organisations there are several factors to consider:

* Data
  + Which attributes of employee data are collected?
  + Does the organisation have engagement survey results/performance metrics, etc. for its employees?
* Model tuning
  + The data balance/imbalance
    - The machine learning algorithms will need to be tuned and/or weighted to best suit the data available.
* Organisational management
  + Management employees need to be prepared to make change and to address dissatisfaction in the workplace.
* Employees
  + Employees need to be open and honest about their workplace experiences so that undesirable matters can be addressed.
    - This can be undertaken anonymously.
* Legal
  + Each country has different employment law(s) that may need consideration.

# 

# Data answer

## Answer

*Which factors (that can be addressed lawfully by an employer) most impact employee turnover?*

This study has demonstrated that those factors that most impact employee turnover are:

* Positive correlations:
  + As the following features increase, attrition increases
    - Over time
    - Business-related travel
* Negative correlations:
  + As the following features increase, attrition rates decrease:
    - Environment Satisfaction
    - Job Satisfaction
    - Job Involvement
    - Years at Company
    - Stock Option Level
    - Years with Current Manager
    - Years in Current Role
    - Job Level
* Neutral corelations
  + The following features had minimal correlation to attrition:
    - Performance Rating
    - Hourly Rate
    - Salary Increase
    - Role
    - Years since last promotion
    - Training
    - Department

For this project, I would proceed with the results from the boosted ensemble modelling, however, depending on other organisational data sets and information the stacked or bagged ensemble models may be more appropriate.

## Confidence

This study, and the three ensemble model results provide a satisfactory outcome. As described above in [Modelling accuracy](#_Modelling_accuracy), a desirable accuracy rate for this modelling is between 80% and 90%; a percentage reached by all three ensemble models.

# 

# Business answer

## Answer

*Can we predict which employees are most likely to leave an organisation?*

From this project, yes, we can predict which employees are likely to leave an organisation using the features described above in [Data answer](#_Data_answer).

## Confidence

With machine learning results of between 80% and 90% accuracy, we can be confident with the answer to the business question.

# Response to stakeholders

As this project has highlighted a number of factors that influence employee attrition, I would make several recommendations to address each of these as detailed below.

**Recommendation 1:**

To address the two factors that most negatively impact attrition:

* Overtime
  + Assess the amount of overtime hours completed and determine whether additional employees are required to address the workload.
  + Assess the priorities of your current employees to determine why over time hours are required.
    - E.g., could timelines for work be extended?
* Business travel
  + With the emergence of more digital ways of working (because of COVID-19 and related restrictions), are there tasks that could be completed remotely?
  + Assessment of the travel and related costs (e.g., fuel, travel allowances, time, etc.) to determine if the return on investment is worth the cost.

**Recommendation 2:**

To address those factors related to employee satisfaction:

* Environment satisfaction
  + Assess which factors contribute to employee satisfaction with their environment:
    - Does the workplace require physical changes? (E.g., a place where people can eat their lunch together).
    - Is the work environment stressful, too noisy, too quiet?
    - Would a hybrid working model with a balance of office hours vs, working from home hours create a more balanced view for employees?
* Job satisfaction
  + Assess the factors that contribute to an employee’s sense of satisfaction with their role:
    - Do they feel that their role is worthwhile to the company?
    - Do they have career progression opportunities?
    - Is the job what they would like to be doing, or is it a temporary position?
      * Could you help them find something they would find more fulfilling within the organisation?

**Recommendation 3:**

There are some clear indicators that stability in the workplace contributes to employee retention. From this project, years at company, years with current manager, and years in current role play a strong part in employee retention.

With this information, I would recommend that an organisation look for ways to provide stability and certainty for their employees. This could be through learning and development initiatives, internal promotions, salary increases (as appropriate), career opportunities within an employee’s current role, etc.

**Recommendation 4:**

To ensure that an organisation reaches the desired level of 10% or less attrition, I would recommend undertaking:

* Annual performance reviews
* Annual salary reviews
* Annual engagement surveys
* Etc.

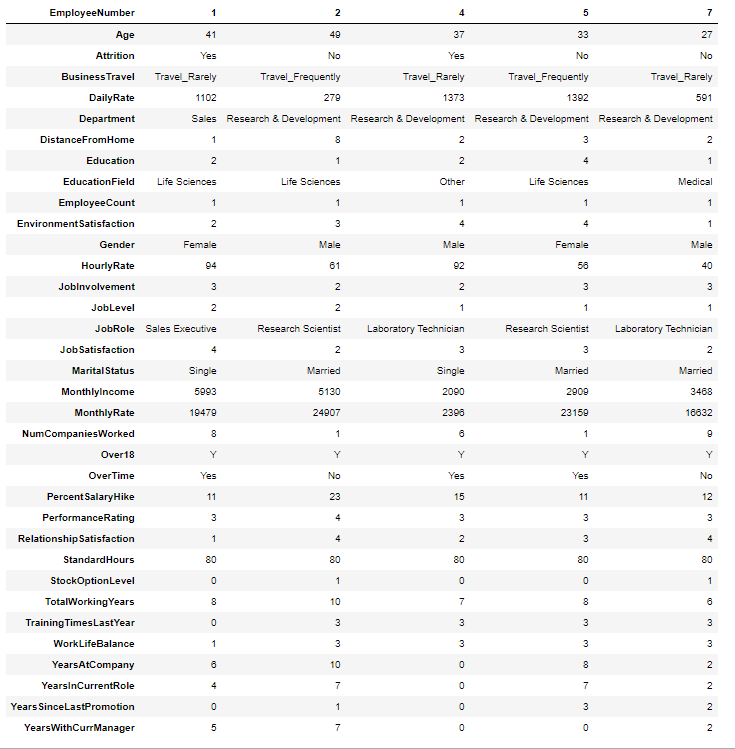
# End-to-end solution

**Figure 6: Diagram of end-to-end solution**



# 

# Appendix 1 – Data snapshot



# Appendix 2 - Data dictionary

Below is a description of the numerical values attributed to features within the data:

|  |  |
| --- | --- |
| **Column name** | **Column details** |
| Education | 1 Below College 2 College 3 Bachelor 4 Master 5 Doctor |
| EnvironmentSatisfaction | 1 Low 2 Medium 3 High 4 Very High |
| JobInvolvement | 1 Low 2 Medium 3 High 4 Very High |
| JobSatisfaction | 1 Low 2 Medium 3 High 4 Very High |
| PerformanceRating | 1 Low 2 Good 3 Excellent 4 Outstanding |
| RelationshipSatisfaction | 1 Low 2 Medium 3 High 4 Very High |
| WorkLifeBalance | 1 Bad 2 Good 3 Better 4 Best |
| Attrition | 0 No  1 Yes |
| BusinessTravel | 0 Non-Travel  1 Travel\_Rarely  2 Travel\_Frequently |
| Department | 1 Sales  2 Research & Development  3 Human Resources |
| JobRole | 1 Sales Executive  2 Research Scientist  3 Laboratory Technician  4 Manufacturing Director  5 Healthcare Representative  6 Manager  7 Sales Representative  8 Research Director  9 Human Resources |
| OverTime | 0 No  1 Yes |

# Appendix 3 - References

**Victorian Public Sector Commission (website)**

* Employee turnover and mobility
  + <https://vpsc.vic.gov.au/data-and-research/data-facts-visuals-state-of-the-sector/employee-turnover-and-mobility/#datasets>
* Employee numbers
  + <https://vpsc.vic.gov.au/data-and-research/data-facts-visuals-state-of-the-sector/employee-numbers/>
* Employee pay and gender pay
  + <https://vpsc.vic.gov.au/data-and-research/data-facts-visuals-state-of-the-sector/employee-pay-and-gender-pay/>

**Articles**

* Chron, Employee Turnover Definitions & Calculations
  + <https://smallbusiness.chron.com/employee-turnover-definitions-calculations-11611.html>
* Chron, What is a Healthy Employee Turnover Rate?
  + <https://smallbusiness.chron.com/healthy-employee-turnover-rate-12145.html>
* Fair Work Ombudsman, Protection from discrimination at work
  + <https://www.fairwork.gov.au/employee-entitlements/protections-at-work/protection-from-discrimination-at-work>
* Forbes, The Cost Of Turnover Can Kill Your Business And Make Things Less Fun
  + <https://www.forbes.com/sites/johnhall/2019/05/09/the-cost-of-turnover-can-kill-your-business-and-make-things-less-fun/?sh=3ce0144d7943>
* Medium, Lima Vallantin, Why you should not trust only in accuracy to measure machine learning performance
  + <https://medium.com/@limavallantin/why-you-should-not-trust-only-in-accuracy-to-measure-machine-learning-performance-a72cf00b4516>
* Workest by Zenefits, What Is Employee Turnover (and Why It Matters)
  + <https://www.zenefits.com/workest/what-is-employee-turnover-and-why-it-matters/>

**Reports**

* Australian Human Resources Institute, Turnover and Retention Research Report,
  + <https://www.ahri.com.au/media/1222/turnover-and-retention-report_final.pdf>

**Coding**

* GitHub
  + <https://github.com/catboost/tutorials/blob/master/model_analysis/shap_values_tutorial.ipynb>
* Kaggle, AdaBoost Classifier Tutorial in Python
  + <https://www.kaggle.com/prashant111/adaboost-classifier-tutorial>
* Kaggle, LightGBM Classifier in Python
  + <https://www.kaggle.com/prashant111/lightgbm-classifier-in-python>
* Towards Data Science, Why you should learn CatBoost now
  + <https://towardsdatascience.com/why-you-should-learn-catboost-now-390fb3895f76>
* Towards Data Science, Ad Demand Forecast with Catboost & LightGBM
  + <https://towardsdatascience.com/ad-demand-forecast-with-catboost-lightgbm-819e5073cd3e>