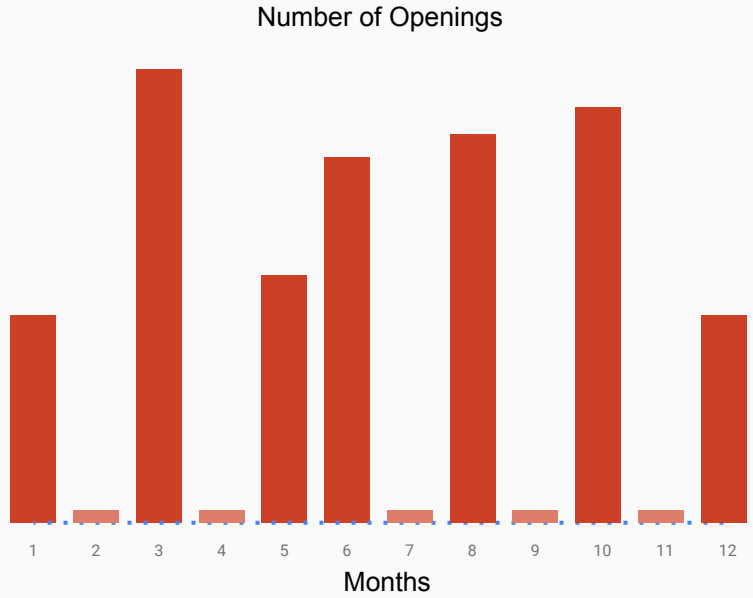
Data Science For People Supply Chain

Siddharth Nishtala, siddharth.nishtala@gmail.com


Mission: To plan supply based on predicted future demand and attrition.

Insights From The Given Dataset

- Hiring in 7 months only
- Each of the 7 months, location and experience grade is the same.
- Total of 222 skill sets were demanded in the last year.
- 61 skill sets appear a maximum of 2 times over the year.



Main Assumptions

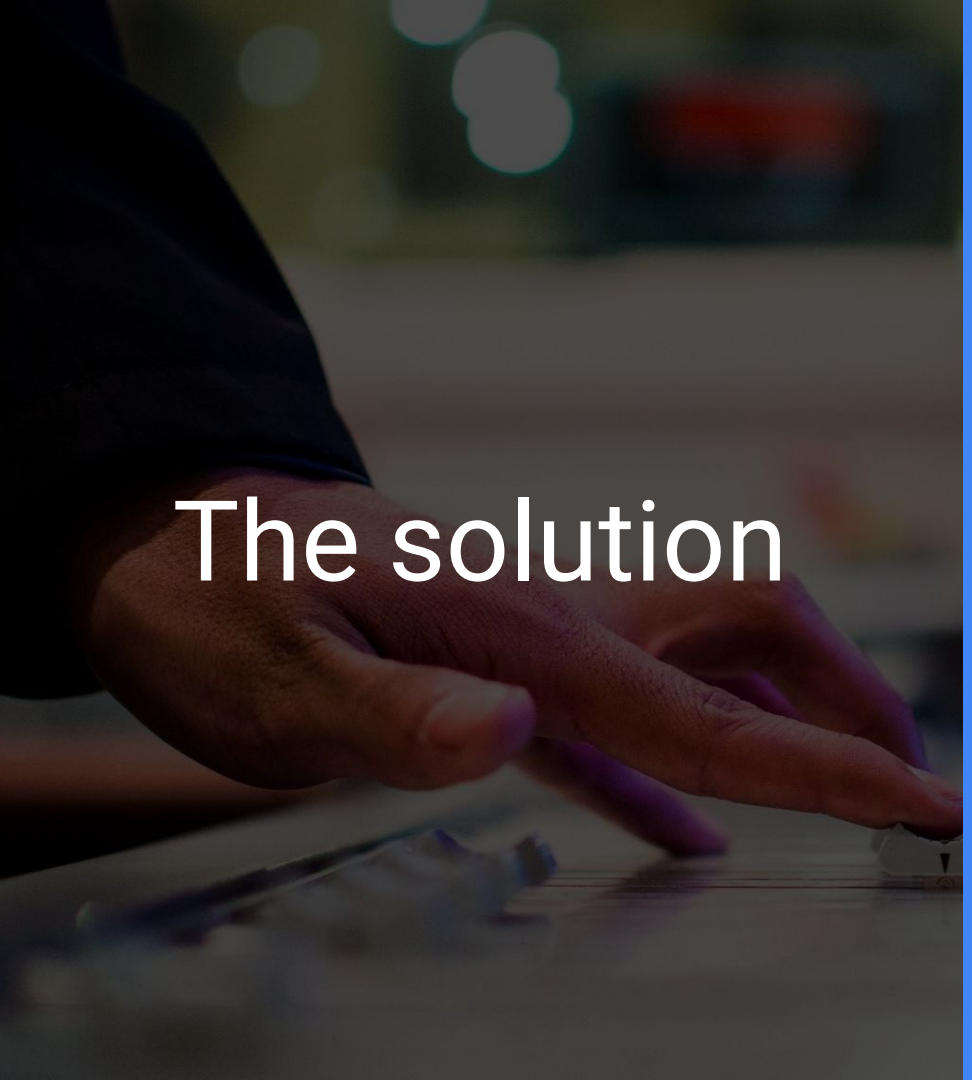
- The aim of the exercise is to **maximise profits** for the company and not the revenue of the company.

  This means that the solution would rather lose some business than spending on an extensive bench.

- Assumptions are taken more to mimic the real life scenario than to maximise model accuracy.

Implementational Assumptions

1. As hiring takes 2 months and no prior information is given about the first two months, it is assumed that **no new employees are taken in the first two months**. This means, the first set of new employees join the company in March.
2. The **given data of last year's demand is missing information** from the months February, April, July, September and November. These values have been interpolated based on techniques discussed in later slides.
3. **Last year's data is representative of a general trend** over the last few years.
4. Notice period of 2 months.



The solution

- ➔ Preprocessing
- ➔ Demand Prediction
- ➔ Supply Planning

Preprocessing

1. Filter out skill sets that occur only once or twice in the entire year. These skill sets introduce noise as we build a demand prediction model. As we filter out skill sets with more occurrences, noise decreases and the overall accuracy of the model increases. Let number of skill sets be ‘n’. (‘n’ = 161)
2. Assign each of the ‘n’ skill sets to numbers from 1 to ‘n’.

Preprocessing

3. Make a matrix ‘E’ of dimensions (12 x ‘n’) with

   E[i , j] = number of employees with skill set j required in month i

4. Make a matrix ‘P’ of dimensions (12 x ‘n’) with

   P[i , j] = E[i , j] / number of employees required in month i

   (P[i , j] = probability of needing an employee with a skill set j in month i)

Preprocessing

5. If ‘i’ is a month with missing data,

   $P[i , j] = (P [i - 1, j] + P [i + 1, j]) / 2$

   $P [i - 1, j]$ is the probability of needing skill set j in the preceding month

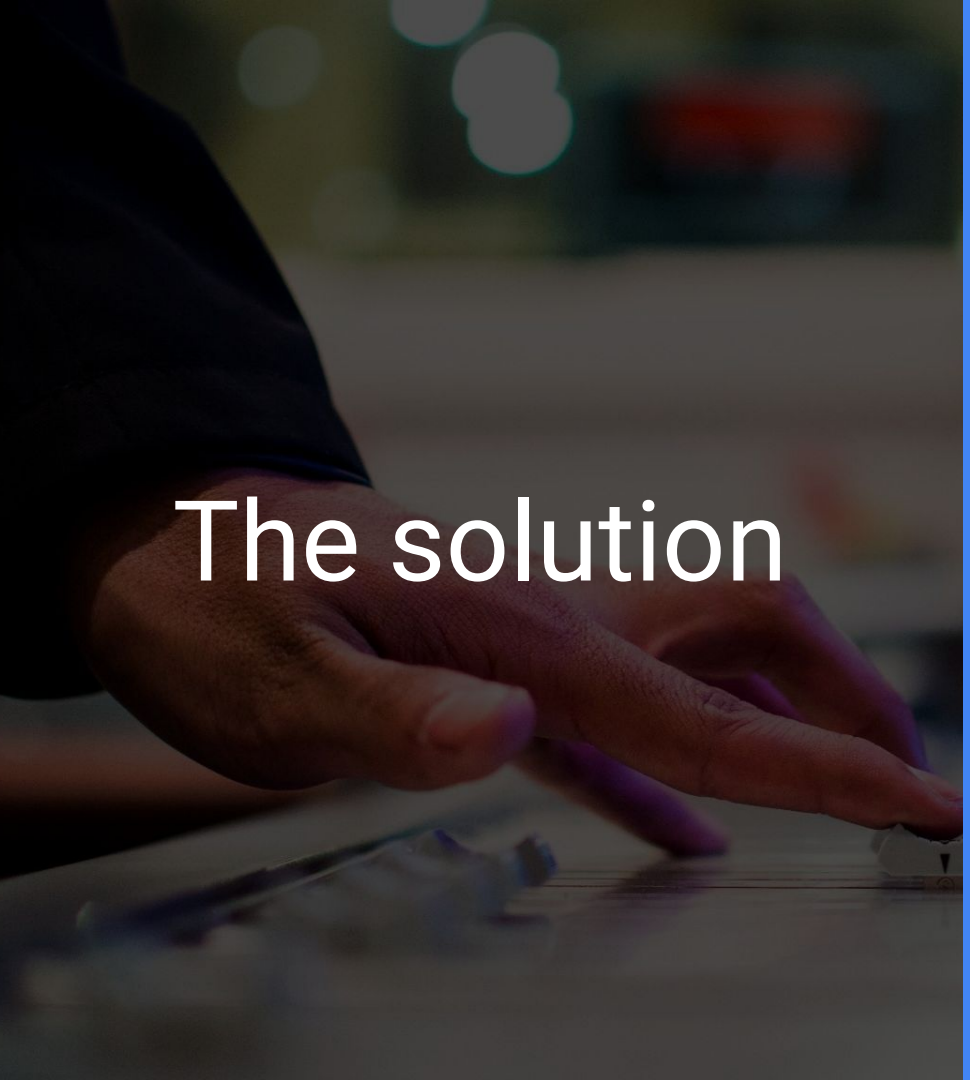
   $P [i + 1, j]$ is the probability of needing skill set j in the next month

   This method is used to deal with missing data.

Preprocessing

6. After adding a random number (as noise) to the monthly demanded employee counts of last year’s data, use the probability matrix ‘P’ to build a more extensive dataset.

   This larger dataset follows the probability distribution of given dataset and can be used to train the demand prediction model.



The solution

- ➔ Preprocessing
- ➔ Demand Prediction
- ➔ Supply Planning

Demand Prediction

The important observation to make while modelling a prediction system is that in the real life scenario, the current demand and the demand of the last few months will be known. This information must be the primary pointer for making predictions of the future demand. Hence, a rolling window regression with a window size of 3 months is used to predict the demand.

Example: January, February and March demand is used to predict the demand of April.

Demand Prediction

Our prediction model is formulated in the following way:

Inputs: Demand of each skillset in month i - 3, month i - 2, month i - 1

Output: Demand of skillset in month i

Input Dimensions: array of length ‘n’ x 3, where ‘n’ is the number of skill sets.

Output Dimension: array of length ‘n’, where ‘n’ is the number of skill sets.

Demand Prediction

The model is to not only predict the skill sets required but also the number of each skill set required.

Neural Network Architecture:

First layer: 750 units with hyperbolic tangent activation, batch normalization and dropout

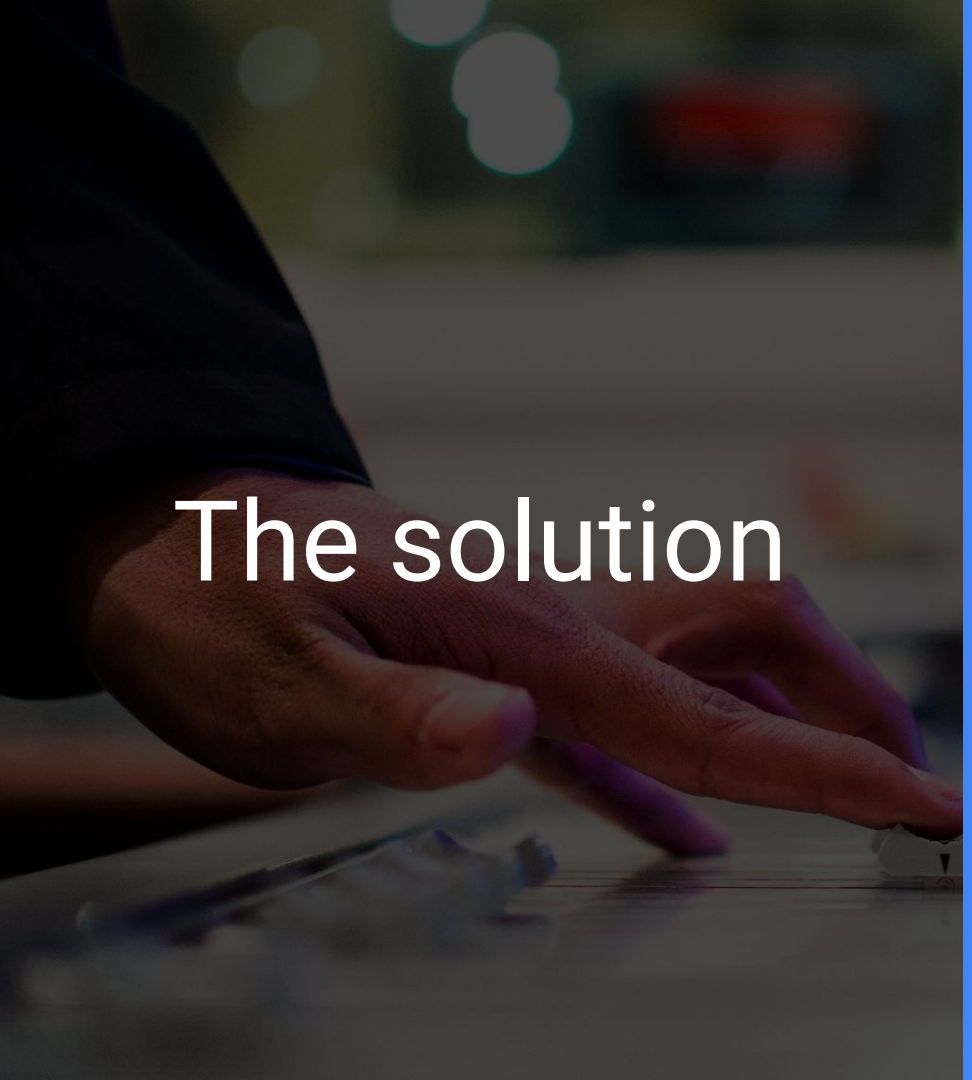
Second layer: 1250 units with hyperbolic tangent activation and batch normalization and dropout

Third layer: 483 units with hyperbolic tangent activation

Output layer: 161 units with no activation

Demand Prediction

In the real life scenario, the current demand will be known. By using the current demand and the demand from the last two months, next month’s demand can be predicted. Now, by using this prediction (i + 1), demand from the current month (i) and demand from the last month (i - 1), the demand for the (i + 2)th month can be predicted. This demand is used for planning the supply.



The solution

- ➔ Preprocessing
- ➔ Demand Prediction
- ➔ Supply Planning

Supply Planning

The inputs to this module per month are:

- Current Demand
- Forecasted Demand
- Previous Billable Employees and Previous Bench
- Billable Employees Resigning and Benched Employees Resigning
- Billable Employees Resigning and Benched Employees Resigning in Future
  This includes the employees resigning in the next two months.
- New Hires

Supply Planning

1. Calculate the current billable employees and current bench.
   current = previous - resigning
2. Add all the new hires to the current bench.
3. Replace the billable resigning employees using the employees from bench.
4. Assign jobs based on current employees from the bench.
5. Calculate the employees that will remain after two months.
   future = current - future resigning

Supply Planning

6. Plan hiring by taking forecasted demand, bench available in (i+2)th month and billable employees leaving in two months.

   Skill sets of employees leaving in two months and skill sets in forecasted demand are needed. If unavailable on the future bench, employees with such skill sets will be hired.

$24,536,895

Total Profit Over 12 Months (averaged over 10 iterations)

Percentage of total business captured: (attrition and demand)	72.71%
Profit per dollar spent:	$0.276962
Total Bench Budget Consumption:	$2488605

86.47%

Training Accuracy

83.97%

Test Accuracy

Demand Prediction Model

Metrics For the Entire Year

January to December (No hiring in Jan-Feb)

Percentage of Total Business Captured (attrition and demand) 72.71%

Percentage of Total Business Captured (attrition) 87.71%

Percentage of Total Business Captured (demand) 65.10%

Metrics After Hiring Begins

March to December

Percentage of Total Business Captured (attrition and demand) 81.18%

Percentage of Total Business Captured (attrition) 96.47%

Percentage of Total Business Captured (demand) 73.69%

NOTE

The current demand used in the simulations was itself the predicted demand of the last three time steps. Ideally, the current demand is observed by the system. As the model was trained to work on a real life scenario, the results obtained from using prediction as observed value are possibly the lower bound of the model's performance. This is because the model could diverge from the pattern over time as the predicted data might not have been a part of the training data.

Thank You.