Theme 2: Intelligent Performance Review Suite

Introduction

Thank you for joining the Intelligent Performance Review Suite webinar, being organized as part of the Radio – EY GDS Hackathon. In this session, we give you a quick overview about the theme, the expectations and answer some of the common questions that you may have.

Business overview

We are trying to analyze the organization capability and consolidate diverse talent data, like structure, performance and potential, leadership pipeline, skill, movements - these are all basic inflammation, syllabus, but in various forms, and in various teams, across the organization. From a solution standpoint, what we are trying to focus on is - how do we integrate and analyze these aspects so that we get better insights regarding this talent pool.

The What - What is the business problem statement

To analyze the organizational capability consolidation of diverse talent data like structure, performance & potential, leadership pipeline, skill, career moves & churn is absolutely critical for the leadership.

Currently, there are multiple tools at EY GDS which provide data or records the input data of employee overall development and they are not integrated due to the above situation, the data doesn't provide any insight / analysis to the business with regard to their talent pool.

The solution being sought is to combine the data from the different sources, intelligently analyse the data and provide better insight with regard to the talent pool.

Dataset overview

So, what kind of data is currently available, or we need to interpret?

- A data with regard to high performance, and potential that is available with the developmental gaps, where we have a construct to analyze that.
- Then we also have skill level, or skill data available for our choice and looking at the developmental plan. That's the analysis that we need.
- There's also data available for attrition analysis - who all are leaving, which location, etc. We're looking at how that helps us to analyze our people movement, how you'll get to recruit for this position.
- There's obviously an annual performance review which flows in from the performance tool that we have.
- And potential, which is based on various conversations and specific behaviors, which the counselor provides us with as inputs to the point, to analyze the potential of the employee.
- And the third one, we're looking at, is, how do we analyze the feedback data coming in from the customers. How is that interpreted and fed into not only the performance, but also to the potential, and any kind of skill gap for measuring that.

There are multiple avenues and sources through which we are getting the employee data. We're also using multiple tools to source the data. There is a plethora of spreadsheets, and mind maps that we do intrinsically just to join the dots and then make an opinion about an employee, whether he or she is a high performer, has the

right potential, skills and competencies to move in the direction based on their aspiration and our requirement. So, we're looking at this from two angles:

1. A data solution - how we synthesize different pieces of data and join the dots together to form an opinion about an employee, specifically around his or her performance potential, and the capabilities or the competencies that he or she predicts.
2. The second aspect is about how all this data can come together in the form of a profile or a capability matrix, based on which we can plot the employee and then work with our leadership for talent reviews to see how the employee can be progressed. Also, what is it that we need to do with the employee in terms of enhancing the right skill set or identifying him or her for a position that is important to the organization.
3. The third and last aspect revolves around the whole inter linkage, and one of the challenges as pointed out earlier, is referring to multiple sources of data. It is important that this tool has, some plugin points through which it can absorb the data coming from our data lake or from any of the sources and help us in putting those dots together.

From a business perspective two things, the data problem, as we call it, and then, the solution of the staging, where all this data can be staged, then, we are able to find some conclusion. The tool or the platform that we're looking at, should be scalable enough, because we are talking about a large-scale organization. It should be customizable in the sense that, if tomorrow we add some more elements or data points as a part of an employee's performance review, then we should be able to scale it up further and integrate newer nuances onto this platform. So, it has to be future ready, both in terms of the architecture, as well as in terms of the technology.

We want to make our decisions based on data, accuracy and speed. While we make accurate decisions, it also has to be made quickly, so that the organization can be benefitted.

Theme details

Problem Statements Details

- Need to build a capability dashboard which will provide the insights to leadership on
  - High Performers and Potential talent with the developmental gaps (in the 15 box construct)
  - Available Skill pool and developmental plan for the talent
  - Churn / Attrition analysis to predict the open positions and internal fulfilment candidate
  - Performance: Align different sources of performance feedback into one tool
  - Potential: Based on conversation and instances of specific behavioural demonstration
  - Perform sentiment analysis on feedbacks / any other customer feedback emails to validate performance and potential

We are looking for a solution where all the data is collated on a single platform and employees can be rated based on different parameters. You can create some data, like a performance tracker and skill tree, in an Excel or any other relevant format. Refer to this data and put your algorithm to that to identify how you can put the place, the employee, in a 15-box grid. So, for this purpose, let us say we are having a feedback tool, where all the employee information and the ratings, and everything will be stored, and an employee feedback is being stored. It will be stored based on different criteria.

- Performance

EmpNo	Provider	Date	Category	Rating	Comments
10001	12345	03-08-2020	Personal	8	
10001	12345	04-08-2020	Team	9	
10001	12345	05-08-2020	Client	7	
11234	13212	06-08-2020	Personal	8	
11234	13212	07-08-2020	Client	8	

As you can see here, there are different parameters - how that person performed at a personal level, team level, and client level, and all these are being tracked against different rating numbers.

Comments also will be coming in from different feedback providers. So, the first thing is to identify how they perform during the last period. With these parameters, you need to identify that. You can use some sentiment analysis, to read these comments, and identify the rating from that, as well, see how better they performed, what is the performance percentage or the number, we can come up for each employee. It should be normalized as well. So that's the reason we are adding provider information, and employee information here. With that number, your algorithm should be able to identify how the employee is rated on the performance level based on other normal sessions and everything.

Potential

EmpNo	Provider	Date	Category	Rating	Comments
10001	23432	03-08-2020	Personal	High	
10001	23432	04-08-2020	Team	Medium	
10001	23432	05-08-2020	Client	Low	
11234	12345	06-08-2020	Personal	High	
11234	12345	07-08-2020	Client	Medium	

And the second parameter added here, is the potential. It will be rated based on some different parameters, like how they are potentially. So, the comments and everything will be also be based on – their potential in doing the things.

The first table will be mainly based on the performance. The comments would be able to register the performances. But, in the second table, it will be mainly depending upon the potential, the growth of the employee or the skill set of that employee in doing the work. You can extrapolate this data by adding more information and make it more meaningful. The comments and everything should be coming based on the performance and potential imposed tables. And based on this, you can get the initial level details.

Skill tracker

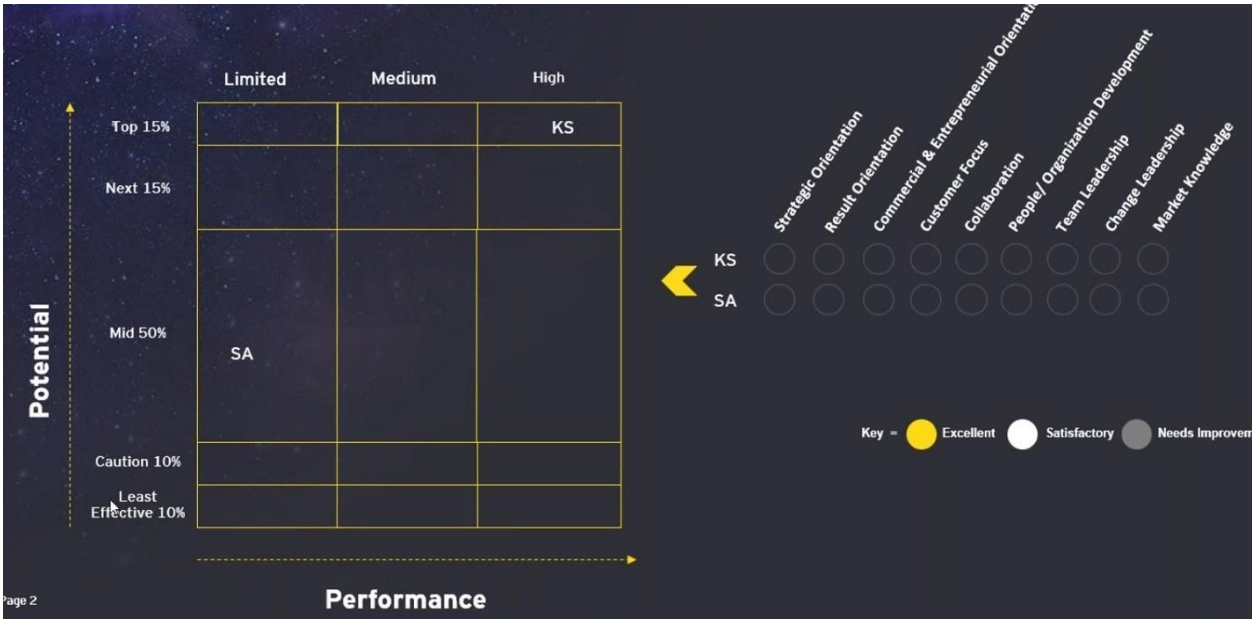
Name	Title	Type	Level	Emp No	Designation Id	Department ID	Country
John	Java	Skilled	4	10001	45	123	India
Mary	Cloud Computing	Certified	4	10002	34	645	China
Rajesh	Negotiation	Skilled	5	10003	54	845	US
Sita	Project Management	Certified	3	10004	23	214	UK

For further information, we are tracking the employees based on the skill tracker. Let's say an employee is getting trained or certified in some new technologies or skills, then these aspects also need to be factored for considering their potential, as they are gaining some additional knowledge and that can be used there.

Performance

Date	Provider	Receiver	Feedback
03-08-2020	12312	11002	I think you did a great job when you ran the all hands meeting. It showed that you are capable of getting people to work together and communicate effectively. I admire your communication skills.

The last data provided here is the general feedback which we are receiving through email, or some other social service. That can also be considered for the performance or the potential. This is the high-level information you will be getting for calculating each employees' potential and the performance. Once you have that data ready, we then need to add that to the below given grid.



You can see potential and performance levels are going on. Based on this, you need to add the employees under 'Limited', 'Medium' or 'High' potential and high performance. Basically, in these 15 boxes we should be able to group all the employees, based on the previous parameters, mentioned above.

Technology stack

1. Solution should ideally suit for deployment in Azure cloud platform:
    - The tools can leverage open source/ custom libraries/ Application Program Interface (API) from cloud platforms.
    - Existing Software As A Service (SaaS)/ hosted solutions (Google Cloud Platform, Amazon Web Service etc.) except Azure should not be leveraged.
2. Preferable technology stack is C#, .net and Python for Artificial Intelligence (AI)/ Machine Learning (ML). You are also free to use libraries that are not licensed under Affero General Public License (AGPL).

It would be good if certain aspects, such as scalability, resilience, etc. can be considered as much as possible. And the solution should be modular enough so that, at some point in time, this particular cloud based solutioning is easily possible.

Q&A

1. **What is the basic input and output that is expected?**

Ans. The input data would be rating received based on personal, team and client aspects. Similarly, we get some inputs on the potential piece. It could be provided by leadership, or there could be some other feedback providers. The different feedbacks on performance and potential that somebody has received. And this includes the verbatim feedback, the skills and the competencies that the employee has.

The output expected is to plot all the employees in 15-box grid and should have an ability to view at function level, business unit level and company level. If you can achieve that, it would be great. Solution should also include the sentiment analysis. Often what we have realized is that the performance inputs, including the potential inputs that we receive sometimes do not tie completely with the verbatim feedback that are being provided through our feedback systems. It could be a simple email from a customer or stakeholder. So, if your algorithm can read or scan through the verbatim feedback that an individual may have received and performance sentiment analysis that tells us whether the statistical performance and potential that is coming up, it matches with this subjective comments or subjective feedback that's received in the form of a text or via email or our feedback systems. It should really point in the same direction, as the performance feedback and the potential feedback, which is more, which is tangible in nature and quantitative in nature.

For example: As the manager of an employee or a team that has 15 employees and based on the input sheets that we just described, I'm getting inputs from different feedback providers – both for my team members' performance. I login into the system as a counselor, I know that I have a team of 15 individuals, and I click on an individual. It should tell me, what are and how he, or she has assessed himself or herself. And then the system should help me in plotting each of these 15 members onto the 15-box performance and potential grid.

2. We have two themes - one around the supply demand that has skill tree and skill level data [AI based capacity management]. The other around customer feedback [Credibility calculator]. Can these two data sources be leveraged to it as well?

Ans. Not really, because, in the first one, when we're talking about an artificial enabled, let us say algorithm to match the supply and demand. Unless the skilled data is provided as one of the inputs from a seller supply side, I do not know exactly what are the attributes that will be provided. If we will define the demand, and what are the attributes that we will provide, redefine the supply. And I believe your engine is then going to do some sort of matching or mapping and then tell us, what is the possible solution. So, if, as a part of the supply attribute, which is the resources available, if the skill tagging for an employee is a part of the data input, yes, for sure, we can use that.

For the second one, the credibility score is not related. Let us say, if there is an employee, who has worked on a particular task and there is feedback that needs to be provided more like what we do, to aggregate services like Uber, or Ola, or even smarter. So, it's not like your regular, and nine-hour task that lasts for the entire year. It's like an open task. It's project-based work that somebody has done, and

we want to provide a comment on how that individual performed, along with some other qualitative comments that yield a credibility score for the next user or subscriber, to view and help in deciding.

3. Do we need to check the match between system feedback and verbal feedback?

Ans. Yes, you can either use sentiment analysis to check or any other relevant method.

4. What is the baseline accuracy currently?

Ans. We would like to believe that it is completely accurate. If you look at the performance tool, we do have a lot of instance-based feedback that will come in for personal attributes versus team and client. And, as you can see, there are ratings as well based on that. Usually, these are stakeholders, my leaders and my customers, who will be giving me this feedback, so, there is that accuracy from an experience perspective. Potential is where accuracy dips a little because performance is based on your KPIs, on what is expected of you, what are your quality metrics, etc. While performance shows me, how have I done so far, potential is about what I can do, or what is the capability that I could have based on some of the behaviors that I'm showing today. Now, that is why the rating is high, medium, and low, and that is based on conversations. Skills, again, are more objective because I have a skill or I don't, and also how skilled I am in terms of my proficiency is something, to a great extent can be accurate because it compares peers.

5. What is the existing automation that is in place?

Ans. We do have existing data in different tools, but do they speak to each other? Nor do we get predictive, or the descriptive analysis out of that. Hence, that's part of the steam, is that, how can we get all of this together, and be able to do some analysis and conviction of the details. We only have spreadsheets, and some data in portals, but nothing that speaks to each other, nothing that's intelligent and nothing that does any sort of data analytics.

6. What is your current data distribution management system? Can you share schema overview for their students?

Ans. Current schema and structure, in terms of how we manage our data today, is mostly through spreadsheets. Different people and teams have different set of spreadsheets. However, there are some common variables in those spreadsheets such as key inputs in terms of skills related to personal, team, client, and then there will be a rating scale onto it. Additionally, different teams will have different data, which may be more variables through which they are measuring, team level performance, or team level potential.

7. Are you expecting a Data Model Proposal as a submission?

Ans. Yes

8. Can you explain who your provider was?

Ans. What we mean by provider, is a feedback provider. It could be anybody in the chain of hierarchy of an individual, or it would be one of our clients, or customers.

9. Is there a correlation expected between feedback and the scale progression of employees?

Ans. Yes, only that we are going to drive a correlation. So, if you guys do that for us, we'll be very happy.

10. Who are the end consumers for this, HR, or any base, any business unit owner?

Ans. Each and every employee receives feedback, irrespective of hierarchy in the organization. It is something that needs to be encouraged. But from an analysis perspective, business units and HR, business unit leaders will come together and take a look at it and see what the trends are, from an overall perspective, how performance is going. However, when you look at the provider and receiver, this is each employee who will get the feedback, who has provided the feedback.

Originally, the idea is, if we are able to build this for the entire organization, both the talent or HR and the business unit leadership will know our capability. In terms of our people, what are the skills? I think somebody talked about skill progression. So, what kind of skill progression we need for our organization to keep on moving forward. It applies both at an individual level, as well as at an organizational level. And the talent and business, they work hand in hand at both the levels.