Original Article

Artificial Intelligence (AI) and Automated Machine Learning Capabilities in SAP Analytics Cloud (SAC)

Sandeep Kumar

SAP Artificial Intelligence and Analytics Leader, Farmers Insurance Group, California, USA.

Corresponding Author : sandeep.96@gmail.com

Abstract - Predictive Analytics is the need of the hour! It has knowingly or unknowingly entered human lives. Many times, one does not even realize how deep is the penetration of AI and its impacts. The shopping list is indexed based on choices and surfing habits. The advertisements are forced based on how much time someone spends viewing a picture or a video. The sprinkler realizes the temperature and adjusts the watering schedule. There are smart kitchens on their way to enter every home in a matter of months.

Organizations have automated their processes for good, and the results have helped them mature and succeed in this brutal competition based on augmented analytics. The organizations are in a rush to identify and implement the latest Artificial Intelligence and Machine learning technologies. This rush has opened the market for creativity and an era of competition towards building and bringing the best-in-class products. It not only undergoes frequent upgrades within itself but fosters an environment of creative construct in the industry.

The heart of any technology to exist and success is dependent on the quality of data that is used to propagate its features. Hence, it brings up a very critical path for organizations to be mindful of what, why and how they are planning to implement Artificial Intelligence and Machine Learning technologies.

Keywords - Business Intelligence, SAP Cloud, Predictive analytics, SAP Analytics Cloud, Data intelligence.

1. Introduction

The last couple of years have been exponentially dramatic for the Information technology industry and equally challenging for companies utilizing reporting, database and planning software. The buzzword has touched every organization and employee, and few are figuring out what this Artificial Intelligence is all about and how one can get there in the shortest span of time. The information technology industry in this space is very crowded. The problem statement is that there is no clear direction on what the best-fit tools or applications one should use to introduce Artificial Intelligence or Machine learning to automate or predict using the enormous amount of data in their organization.

There are several applications in the market; however, especially for companies who are using enterprise resource package software to run the backbone and on the frontlines, example Sales and Distribution, Manufacturing, Finance, Human Resources, Supply Chain, and Project systems can very well adapt to the Artificial Intelligence and Machine learning applications SAP has introduced.

SAP (Systems Applications and Products) is one of the giants in the information technology industry, and they released SAP Analytics Cloud (SAC) as the latest and greatest Artificial Intelligence (AI) powered augmented analytics application. The SAP Analytics Cloud, aka SAC, is the new generation of software-as-a-service cloud application solutions that has the capability to be utilized by organizations for augmented analytics.

The product is a new generation of cloud offering from SAP, the industry leader for the last 40 years in the enterprise resource planning software industry and is known in the market for giving an unmatched business warehouse system.

There are several tools and applications having functionalities both out of the box and custom that provide similar solutions and models; however, if the backbone of the company's data warehouse is SAP, then there is not anything better than an SAP application having a native connection and integration like SAP Analytics Cloud.

2. Capabilities

2.1. What are the Capabilities of SAP Analytics Cloud in Predictive Modeling and Why?

The SAP Analytics Cloud comes with inbuilt machine
The components of the SAP Analytics Cloud that enable predictive capabilities are the following:

- Machine Learning: The automated learning of processes based on the data.
- Artificial Intelligence: It is the ability of a machine to understand and mimic actions based on human perception and feelings. These are generally a combination of codes, multilevel algorithms and process techniques that result in achieving intelligence.
- Augmented Analytics: Analytics based on continuous learning and NLPs aka natural language processing, to perform specific actions such as human cognition and behavior, which could be improved to be better than a human.

2.2. SAP Analytics Cloud – Smart Features

SAP Analytics cloud also has inbuilt smart features that come out of the box and can be used by business users in self-service capability. This was aimed at increasing the tool's adoption in the user community. The smart features are:

- Search to Insight: All one must do is type a question.
  It is a feature to build charts within SAP Analytics Cloud based on the input and discover relationships within data to present the information most aptly. One can use natural language in the search box to see the results. A simple example would be to search for "Company XXX Revenue by current year vs Previous year". The inbuilt algorithm builds a chart to be analyzed and used further in the SAP Analytics Cloud stories/dashboards. This feature is of great help to the superusers.

- Smart Discovery: This feature uses a strong combination of statistical analysis and predictive analytics and helps with data exploration using machine learning capabilities. An entire story can be built by this feature if data, dimensions and metrics are provided. Data and reports can be analyzed at a granular level to narrow down the requirement.

- Smart Insight: This feature lets the user look deeper into the insights with further support to identify criteria and key influencing factors for a pattern. The algorithm further drills down based on the data and dimensions to provide more content around the pattern. Smart insight identified in the panel can be further drilled down for deep insights to the lowest grain, explaining that one can add to the story.

learning and artificial intelligence capability for implementing predictive analytics on the business data. It empowers the business community with access to deep insights and enables faster data-driven decision-making

- Smart Predict: The built-in feature helps businesses predict outcomes based on the data supplied and by training the model. This way, businesses can spend more time brainstorming what could be predicted and the number of years of data rather than spending time on how to attain predictions. This feature totally transforms the time the business spends in attaining and analyzing the predictions and replaces all the local data manipulations in Excel with much-sophisticated machine learning-driven predictions.

There are 3 out of the box prediction models provided:

- Classification model– Used for determining probabilities of an event.

- Regression model– Used to identify numbers and their influencers.

- Time series forecast model– Used to forecast based on key dimensions.

The smart feature is the next level of augmented analytics to enhance the capability of the business community to utilize the best of both worlds, i.e., machine learning analytics plus human creativity.

3. Smart Predict Application in Building Predictive Scenarios

Smart prediction of data is configured using a predictive scenario workspace. This is a preconfigured workspace, and it is used to create multiple models to address business enquiries, which require prediction of probabilities, numerical data and its influencers, and forecasts based on the historical dataset.
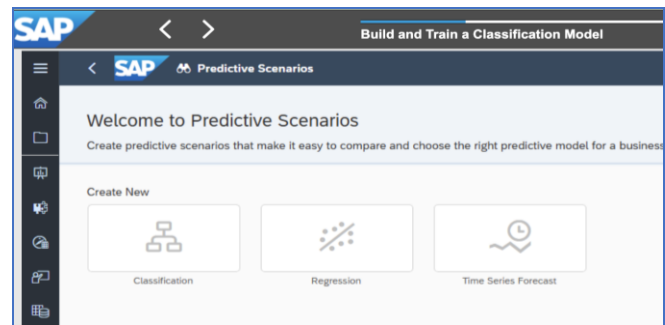

Fig. 1 SAP SAC Predictive scenario workspace

Now, let us look at what is required to launch a predictive model. There are 3 phases in which the process can be built and applied.

- Source data for building the predictive model– The input dataset provided is the core of the entire predictive analytics process and its outcome. The predictive model is trained based on the actual results from the historical data. The artificial intelligence and machine learning capability identifies patterns/influencers in the historical data that resulted in the known outcome.
- Target data is the data on which the learnings from the predictive model are to be applied to predict the probabilities.

- Target outcome- This results from the machine learning run on the target data. The results are predicted probabilities that can be saved in a file or on the SAP database (HANA)*. The probabilities will be as great as the input, i.e., source data and is pivotal for business in making data-driven decisions.
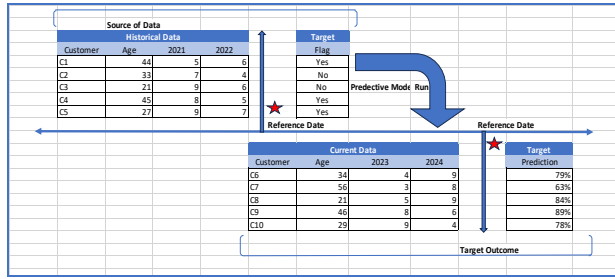

Fig. 2 SAP SAC Predictive modeling process

4. Why does SAP Analytics Cloud no SAP Datawarehouse have an Edge Over other Applications?

- Self-Service Capability: SAP is a strong promoter towards business user drive and runs the SAP Analytics Cloud with IT support. The modeling and rendering options are designed to keep business users first.

- Integration: Native Integration with all the underlying systems helps maintain data integration, accuracy, and completeness.

- Live Connection: The user has the option to run the predictive on data read from live systems, which can be hosted in the SAP Analytics Cloud in near real-time. Major advantages compared to other sources, which this is a limitation.

- Security: Fully supports and utilizes security roles and authorization built into the SAP source system.

- Smart Features: Out-of-the-box supported smart features for data discovery, natural language model providing research insights and predictive analytics.

- Reduced dependence on IT: Low code environment and functionality empowering users to build their model and analytics.

5. Conclusion

The race for organizations to be the best in various domains like Sales, Finance, Human Resources, Marketing, etc., has created a space for niche technologies like Artificial Intelligence and Machine Learning. This space has a lot of competition today; however, SAP Analytics Cloud has emerged as a leader and has the capability and features to run multi-scenario predictive modeling and augmented analytics. It's state-of-the-art, out-of-the-box capabilities and extensions to build custom applications can help organizations emerge as a winner in the future.

Predictive Analytics Components
1. Story: Reports/dashboard equivalent
2. Variable: It is a column in the dataset
3. Target: Placeholder for the variable value to be predicted 7
4. Influencers: Variables which describe the data
5. Hold Out Samples: Sample of observations learnt during the model's training.
6. Training/Validation Subset: Input data is split into training and validation subsets. For classification and regression predictive models, the ratio is 75% vs 25% based on random records. For Timeseries, it is 75% vs 25% based on the sequence of records.
7. Variable Types: Statistical, Data type and Role
8. Statistical Types: Normal, ordinal and Continuous variables
a. Normal Variable: Discrete but unordered set of values
b. Ordinal Variables: Discrete but ordered sets of values.
c. Continuous Variables: Numerical that can take values including fractions and decimal places.
9. Data Types: Two data types exist- Quantitative or numerical data and Qualitative or categorical data.
10. Role: Variable roles are of two types,
a. Target or Signal variable- A variable being predicted, and there can be only one such variable.
b. Date Variable- Variable used for date dimensions. It is mandatory to have one date variable.
c. Entity- is equivalent to a dimension in data.
11. Storage Format: String, Integer, Number, Boolean, Date, Time, Date & Time
12. Global Performance Indicator:
a. Predictive Power- It is a measure of the accuracy of the model and should be close to 100%
b. Predictive confidence- It is the ability of the predictive model to achieve the same grade of accuracy when applied to a new dataset of the same type. It should be close to 100%
13. Confusion Matrix: It is also known as error matrix.

14. Sensitivity: It is the probability of the model to predict a positive outcome when the outcome is positive
15. Specificity: it is the probability of the model to predict a negative outcome when the outcome is negative
16. Multiple Linear Regression: This explains the relationship between continuous target variables and two or more influencer variables.
17. Root Mean Squared Error: It is the difference between predicted values vs actual values.
18. Forecast Horizon: It is the future period for which a Timeseries forecast model needs to forecast. The standard ratio is 5:1, which means 5 years of data for predicting 1 year of the future.
19. Expected MAPE: It is used to check the quality of the prediction of a predictive model. It is the mean of all calculated horizon MAPE values.
20. HANA-   SAP HANA is an in-memory, column-oriented, relational database management system developed and marketed by SAP.

References

[1] SAP. [Online]. Available: https://www.sap.com/products/technology-platform/cloud-analytics.html

[2] Ned_Falk. [Online]. Available: https://sapinsider.org/an-introduction-to-sap-analytics-cloud-predictive-analytics/

[3] Seth_Magehee. [Online]. Available: https://blogs.sap.com/2022/06/01/sap-analytics-cloud-introduction-to-predictive-scenarios-and-applications/

[4] Help.sap.com. [Online]. Available: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD?version=release

[5] Thierry_Brunet. [Online]. Available: https://blogs.sap.com/2020/06/19/classification-in-sap-analytics-cloud-in-detail/

[6] Thierry_Brunet. [Online]. Available: https://blogs.sap.com/2020/08/13/regression-in-sap-analytics-cloud-in-detail/

[7] Thierry_Brunet. [Online]. Available: https://blogs.sap.com/2019/10/07/time-series-forecasting-in-sap-analytics-cloud-in-detail/