AA - API DataRobot Deployment

The data pipeline offers a self-service framework for building, configuring and utilizing individual components for moving data between systems. As such, this enables a seamless scoring process for the Advanced Analytics predictive models which further streamlines the model deployment phase of the data science value chain.

Data pipeline Configurations

The table below provides the required configuration elements that drive the tailored data pipeline solution for Advanced Analytics.

Config Component	Config Name
Tenant	Advanced Analytics
Tenant Projects	The following projects have been created as part of the testing and future embeddings: • **ws_ei_movements_wb -** Facilitating the AM movements write-back. • **ws_sil_lapse_reten -** Facilitates the scoring pattern for the SIL early lapse model. • **ws_s14_pred_reten -** Facilitates the scoring pattern for the S14 predictive model.
Hive service account	svcDevSbiAaDpNfClAcc
SAP HANA service account	svcDevSbiAaDpHaAcc

Data pipeline Considerations

The table below provides some of the considerations that need to be incorporated to cater for some of the existing and future AA use cases based on the testing conducted so far:

Component	Description
quote delimiters	Upon testing for the SIL and S14 use cases there have been errors as a result of the was the delimiter splits have been defined. This in some instances creates more fields which inconsistent datatypes. The following script serves as an example which relates to S14 prototype use-case: AA S14 - Score_Insert_component.txt
Hard referencing hive field name	There exists a limitation within the hive statement query of the "dr" component which retains the table.column name header as the actual column name if a "*select **" is used within the query. A work-around has been created to hard-code the fields required for the scoring. This is only temporary and should be amended in future releases.
Dev instance API	This component refers to an API scoring process which involves the Non-ML-Ops server. Given that DataRobot has limited slots for deployments within the ML Ops environment, a new sub-pattern is required to be established in order to facilitate this process. As it stands, majority of the development work has been complete, however this requires overall testing by the AA modeler. The following Jira presents the request : ☑ ~~GBIP-60635~~ - Incorporate new API scoring pattern as part of the data pipeline for AA team `DONE`
Write-back configuration VNB	This refers to the establishment of a formal write-back of the VNB information from Actuarial reporting using the data-pipeline solution. This can be initiated with the ingestion team to support and maintain the write-back. It must be noted that the following two configurations must be taken into account for the formal write-back of this data: • Chunking mechanism attributes: SCENARIO_CD and FIELD_NM. This enables a more efficient partitioning for the write-back. • The related analytical privilege that has been set up is: bi.security.data.spf.actuarial.actval.external:AP_ACTVAL_AA_ADV. analyticprivilege. - This will allow for service account to have access to the view of choice.