

Insurance Loss Estimation using Remote Sensing data

Introduction

With the increase in the number of satellite launches and the dramatic improvements in sensing technology (e.g., better cameras with hyper-spectral imaging, synthetic aperture radar, and lidar), remote sensing data now has much better quality in resolution and spatio-temporal coverage. There is demand for using this data to solve business issues across a wide range of industries, including insurance, urban planning, agriculture, climate change, and flood prevention. This creates a need for ML/AI technology solutions for remote sensing data.

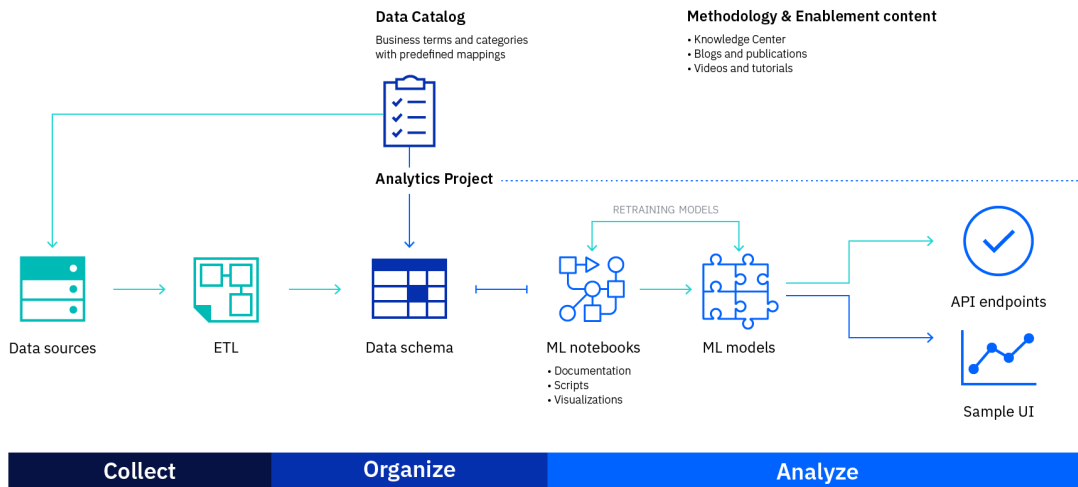
Insurance Loss Estimation Using Remote Sensing Data Industry Accelerator shows how to derive insights from remote sensing data, by utilizing an example of studying flooding events for assisting insurance claims. The idea in this accelerator is to study satellite images before and after a certain event, and with the help of spatiotemporal analysis and ML/AI techniques, we can get insights into which regions are affected and to what extent. Such insights provide valuable information to insurance companies on property damage, allowing claims to be processed more efficiently.

For this specific usecase, we picked a category 4 storm event -- Hurricane Laura -- which caused damage to many areas of Louisiana, United States in August, 2020. Insurance companies can use this accelerator to understand remote sensing capabilities to study the flood event, get insights on impacted regions and then predict estimated claims for individual properties.

Your data scientists can use the sample notebooks, remote sensing analysis, predictive model, and dashboard to accelerate data preparation, machine learning modelling, and data reporting. All the functionality in this accelerator can be generalized and used towards other use cases for remote sensing data.

What you provide

What we provide via Industry Accelerators



- [Instructions](#)
- [Sample data assets](#)
- [Notebooks](#)
- [R Shiny dashboard](#)
- [Sample business glossary](#)

Instructions

Follow these steps to implement the industry accelerator:

1. Navigate to the **Assets** tab and scroll to the **Notebooks** section.
2. Edit the **1-impact-region-analysis-with-remote-sensing** notebook by clicking the edit icon that looks like a tiny pencil next to the notebook name. This notebook analyses impacted regions and generates impact scores for all the regions. Follow the instructions in the notebook to step through the execution. Alternatively, you can run the notebook from the Jobs tab by executing **1-impact-region-analysis-with-remote-sensing-notebook-job**.
3. Edit and run the **2-model-training-and-deployments** notebook. This notebook prepares the data, builds ML models, and deploys the model. It also deploys the r-shiny dashboard and generates a URL to launch the r-shiny dashboard. Alternatively, you can run the notebook from the Jobs tab by executing **2-model-training-and-deployments-notebook-job**.
4. Launch the r-shiny dashboard from one of the following ways.
 - Open the URL generated by deploying r-shiny dashboard in **2-model-training-and-deployments** notebook.
 - Navigate to **Deployments -> Spaces -> Insurance Loss Estimation Using Remote Sensing Space -> Deployments -> Insurance-Loss-Estimation-Using-Remote-Sensing-Shiny-App** to find the URL of deployed r-shiny dashboard and open it in a new tab.
 - Run the dashboard from RStudio console by completing these steps:
 - i. Download the Insurance-Loss-Estimation-Using-Remote-Sensing-

- Dashboard .zip file from the Data assets section of the **Assets** page. If you don't see the file, click **View All** to display the full list of assets.
- ii. Click **Launch IDE > RStudio** on the menu bar.
 - iii. In the **Files** pane, select the **Upload** toolbar button and upload the Insurance-Loss-Estimation-Using-Remote-Sensing-Dashboard .zip file into RStudio.
 - iv. Select the app .R file, and click the **Run App** toolbar button to launch the dashboard. If you see a warning message that certain packages are not installed, you can ignore it because the packages will be installed first time you run the app.
5. Once the app has launched, you can perform model scoring in real time by entering your username and password on the **Property View** tab.
 6. Optional. To connect the data assets used in this accelerator to the business terms in Watson Knowledge Catalog, you can edit and run the **0-map-business-terms-to-data-headers** notebook. Enter the authentication details required in the first few cells.

Sample data assets

These sample data files that act as dimensional and fact tables are included in the project on the **Assets** page:

Remote Sensing Input.csv: Property Ids and property coordinates in Cameron town, Louisiana.

Insurance Loss Claims.csv: Customer's property information including property value, construction details, size, property damage amount etc. and insurance information including sum insured limit, previous claims, estimated insurance claim etc.

b03_before.tif: Green band satellite image of Cameron town captured before Hurricane Laura.

b03_after.tif: Green band satellite image of Cameron town captured after Hurricane Laura.

b08_before.tif: Near infrared band satellite image of Cameron town captured before Hurricane Laura.

b08_after.tif: Near infrared band satellite image of Cameron town captured after Hurricane Laura.

Additionally, there are other 2 datasets created via the analytics project:

Remote Sensing Output.csv: Flood impact details on the properties provided in Remote Sensing Input.csv. This dataset is generated by the notebook 1-Impact-Region-Analysis-with-Remote-Sensing.

model_output_summary.csv: Consolidated prepped data after combining remote sensing and insurance loss claims datasets for exploratory data analysis and data visualization in the R shiny dashboard.

Notebooks

Follow the instructions in the notebooks to step through the execution.

- **1-impact-region-analysis-with-remote-sensing:** This notebook performs the following functions:

- Study satellite images before and after a certain event (e.g. hurricane Laura) with the help of spatiotemporal analysis and ML/AI techniques
- Get insights to understand which regions are affected and to what extent
- Read in property coordinates to calculate impact scores on each property and save out the result.
- **2-model-training-and-deployments:** This notebook performs the following functions:
 - Load Remote Sensing Output and Insurance Loss Claims
 - Prepare and clean data for model training
 - Analyze correlations
 - Build ML models, Analyze and visualize the data
 - Select best performing ML model and save to Cloud Pak for Data
 - Create a deployment space and store the pipeline in the space and deploy
 - Store and deploy R Shiny app
 - Generate URL to view the app.
- **0-map-business-terms-to-data-headers:** This optional notebook performs the following functions:
 - Publish the **Insurance Loss Claims.csv** and **Remote Sensing Input.csv** files into a specified catalog.
 - Read mappings from **insurance-loss-estimation-using-remote-sensing-data-map-terms.csv** and applies business terms to the published dataset headers.

Jobs

Navigate to **Jobs** tab to execute following jobs.

1-impact-region-analysis-with-remote-sensing-notebook-job: Runs the 1-impact-region-analysis-with-remote-sensing notebook end to end.

2-model-training-and-deployments-notebook-job: Runs the 2-model-training-and-deployments notebook end to end.

R Shiny dashboard

The R Shiny dashboard displays flood impact insights on Cameron town, property details and scores new data. The dashboard has the following tabs:

- **Remote Sensing:** Displays the impact zones and properties of Cameron town. By clicking on the impact zones in the map the dashboard displays properties and insurance details specific to the impact zone. The tab also shows the satellite images before and after the storm event.
- **Property View:** Targets individual property and insurance information, displays an impact score calculated by the remote sensing exercise. It provides the option to run the model webservice which predicts the estimated claim amount for the property.
- **Simulation Tool :** This tab contains a form with all model inputs. The user can change any of these inputs and see the impact that the change has on predicting the estimated claim amount.

Business glossary for use with Watson Knowledge Catalog

Optionally, you can import the glossary of business terms into Watson Knowledge Catalog to get started on data governance using the below files available in the project tar file.

The `insurance-loss-estimation-using-remote-sensing-data-glossary-categories.csv` file defines the main and sub categories for the business terms.

The `insurance-loss-estimation-using-remote-sensing-data-glossary-terms.csv` file defines the business terms, category of the business terms and their Related Terms/Part of Terms, if applicable.

Once the glossary is imported into Watson Knowledge Catalog, **Publish** the Business terms, Navigate to **Governance > Categories > Industry Accelerator > Insurance Loss Estimation Using Remote Sensing Data** to explore the glossary contents.

Terms and Conditions

This project contains Sample Materials, provided under license.

Licensed Materials - Property of IBM.

© Copyright IBM Corp. 2020, 2021. All Rights Reserved.

US Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.