Predictive Testing Strategy at BMW Group

Using the Deep Learning Toolkit for Splunk

Andreas Schoch

Machine Learning Researcher | IT Innovation Lab, BMW Group

Philipp Drieger

Principal Machine Learning Architect | Splunk

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or plans of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results may differ materially. The forward-looking statements made in the this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, it may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements made herein.

In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only, and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionalities described or to include any such feature or functionality in a future release.

Andreas Schoch

Machine Learning Researcher | IT Innovation Lab, BMW Group

Philipp Drieger

Principal Machine Learning Architect | Splunk

Agenda

BMW Innovation Lab
– Who we are and what we do

Use Case and Business Objective
– What is the challenge we want to tackle

Data Science Solution
– How we use the Deep Learning Toolkit for Splunk

Future Plans
– What's next

splunk> .conf20

BMW Innovation Lab

“We like to innovate, we like to do things that haven‘t been done before, and we like to be leading-edge in technologies.”

“The Innovation Lab can be thought of as an interface between the present and the future quality management, where concepts can be integrated into series production.”

BMW Innovation Lab and Splunk

5+ years of collaboration and innovation

Our journey of innovations with Splunk:

- conf15: Keynote Presentation
- conf16: Save Energy with Splunk
- conf17: The Next Level of Quality Assurance at BMW with the Machine Learning Toolkit
- conf18: Keynote Presentation for Machine Learning
- conf19: Image Indexing Framework for Image Search and Deep Learning Applications
- conf20: Predictive Testing Strategy at BMW Group Using the Deep Learning Toolkit for Splunk

Use Case and Business Objectives

Predictive Testing Strategy at BMW Group

“Everything we do we believe in challenging the status quo.”

splunk> .conf20

"The way we challenge the status quo is by **creating individual, dynamic and proactive test processes** in order to increase process efficiency."

splunk> .conf20

“Can we predict possible error patterns based on a car’s specific product configuration?”

Theoretical Approach

Need for predictive testing strategy in order to increase process efficiency

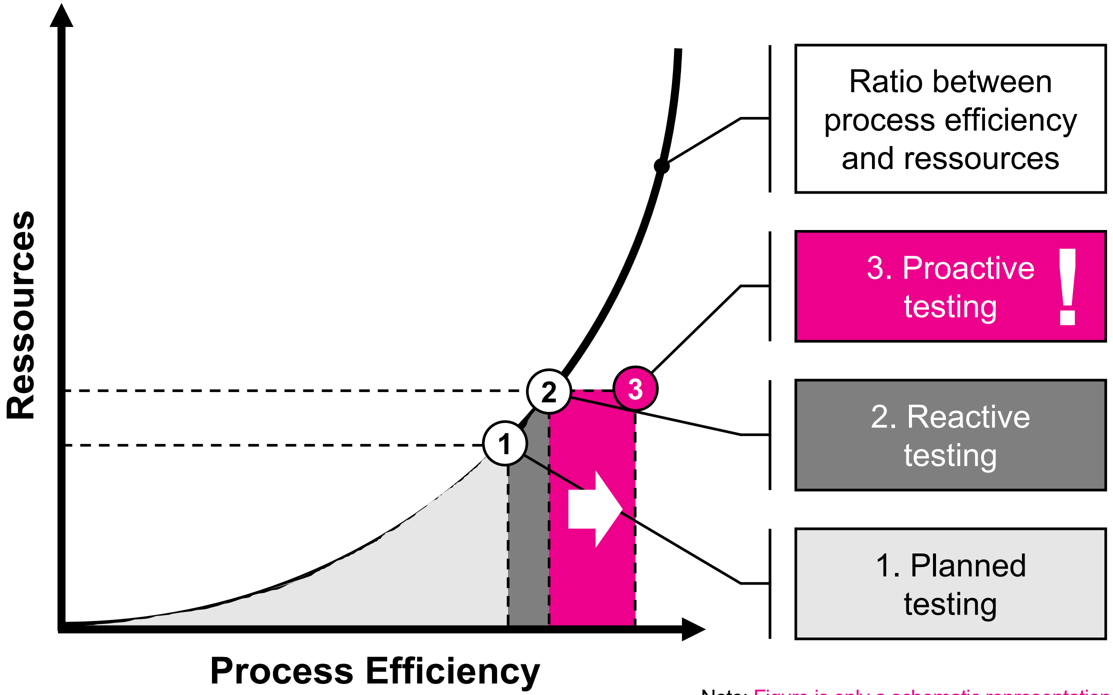
Status Quo

- High degree of process efficiency due to planned and reactive testing (stage 1 and 2)
- Implemented test processes are capable of detecting all relevant error patterns during production
- Extremely high degree of process efficiency and vehicle protection

Challenge

- Implement a predictive testing strategy, which is...
  - ...individual = create vehicle-specific testing proposals
  - ...dynamic = update testing proposals
  - ...proactive = take concrete actions to avoid possible errors

  ...in order to increase process efficiency without any additional resources (tackle stage 3)



Note: Figure is only a schematic representation.

Practical Implementation

Implement dynamic quality gate to spotlight and prioritize possible error patterns

High-level process description

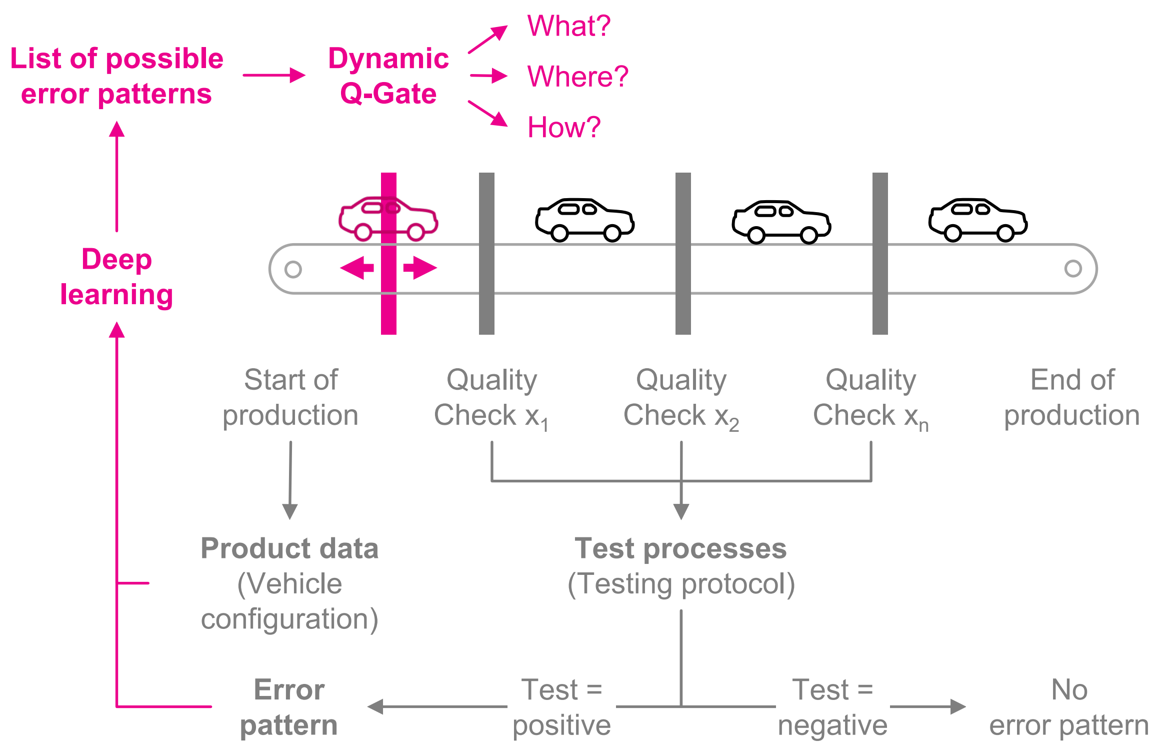
- Lots of standard quality checks (x_n) through production
- All testing results (+/-) are splunked in detail

Approach

- Combine product data and deep learning to predict possible error patterns on vehicle level
- Implement a dynamic quality gate at production based on a vehicle-specific list of possible error patterns

Objective

- What: prioritize possible error patterns
- Where: detect error patterns where they are caused
- How: recommend qualified testing activities



Digital Transformation

Increase process efficiency using Deep Learning Toolkit for Splunk

Now

"Splunk Enterprise provides an integrated view on the vehicle's data and enables production teams to keep track of vehicles in order to assure highest production quality."

Future

"Deep Learning Toolkit for Splunk is the brain behind dynamic and proactive test processes and enables production teams to make smarter decision in order to increase process efficiency."

Data Science Solution

Using the Deep Learning Toolkit for Splunk

What is DLTK?

Freely available app for advanced data science projects using any open source AI frameworks

Speed up your data science projects with GPU accelerated containers

Seamlessly integrate & operationalize with Splunk Enterprise

Download the app from splunkbase:
https://splunkbase.splunk.com/app/4607/

Accelerate Your Data Science Innovations

Leverage GPU powered model building and easily integrate with Splunk

Technical setup used in this project:

- NVIDIA DGX Workstation with 4 GPUs
- Splunk Enterprise 8.0
- Deep Learning Toolkit for Splunk 3.2
- Tensorflow based Deep Neural Network
- Historical Data of 100,000+ Produced Vehicles

splunk® > turn data into doing™

Image Source for NVIDIA DGX Workstation: https://www.nvidia.com/en-us/data-center/dgx-station/

Data Science Approach

Following five typical steps for an applied data science project

Frame the Problem

Predict possible error patterns based on vehicle data

Get Data and Explore

Product and process data of 100K+ vehicles with high feature cardinality

Create Models

Use the Deep Learning Toolkit for Splunk to create neural network models

Evaluate and Understand

Evaluate, Understand and Explain Model Predictions within Business Context

Present and Operationalize

Predictive Testing Dashboard and Operationalize for daily Reporting

Frame the Problem

Can we predict possible error patterns based on a car's specific product configuration?

Use Case Description

Predict error patterns based on vehicle data

Concept

- Use product data (vehicle configuration) and specific error patterns
- Learn the connection between product data and error patterns
- Try out different machine learning concepts and approaches

Results and outcomes

- Find the best model to predict possible error patterns
- Shortlist vehicles according to their highest testing probability
- Provide a final dashboard for production teams

Note: All data used in the presentation are synthetically produced data.

Get Data and Explore

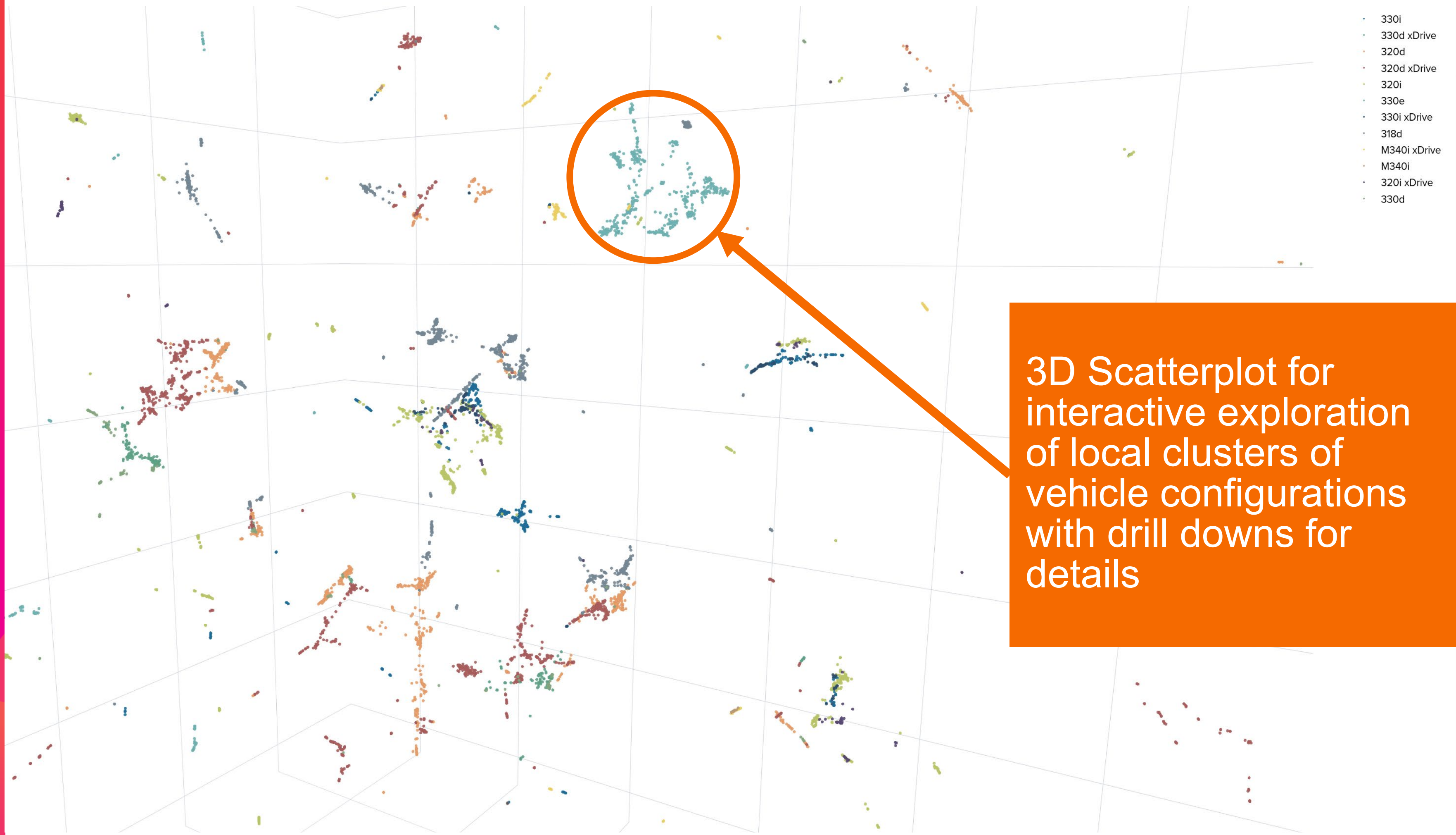
UMAP based exploration of high dimensional dataset

+

Apply Clustering and detect groups with increased frequency of error combinations

Dimensionality Reduction and Exploration of the Feature Space

Compute meaningful features and use for visualizations



3D Scatterplot for interactive exploration of local clusters of vehicle configurations with drill downs for details

More information on UMAP: https://umap-learn.readthedocs.io/en/latest/

Note: All data used in the presentation are synthetically produced data.

Create Models

Baseline with a basic Random Forest Classifier

Create a Baseline Model

Using Machine Learning Toolkit's RandomForestClassifier

Algorithm: RandomForestClassifier
Field to predict: testing_true
Fields to use for predicting: date_wday, dhour_... (15)
Split for training / test: **no split**

N Estimators: 100
Max Depth: (optional)
Max Features: (optional)
Min Samples Split: (optional)
Max Leaf Nodes: (optional)

Notes: (optional)

Fit Model | Open in Search | Show SPL

How does the model perform on all training data?

Prediction Results

testing_true	predicted(testing_true)	date_wday	dhour_periode	model	description	drive	hybrid	motor	country	x_1	x_2	motor_family	motor_form	motor_modul	motor_range_lifecycle	motor_type
1	1	wednesday	11-4	G21	320d xDrive	LL	HYBR	XD5	FR	52	31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1	wednesday	10-4	G21	330i xDrive	LL	[illegible]	B48	PL	63	26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1	wednesday	10-3	G21	320d xDrive	LL	HYBR	XD5	FR	52	31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1	wednesday	9-1	G21	320d xDrive	[illegible]	HYBR	XD5	FR	52	31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
0	1	wednesday	9-1	G20	M340i xDrive	LL	NOHY	B58	CH	71	35	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1	wednesday	8-1	G21	320d xDr[illegible]e	LL	HYBR	XD5	DE	66	34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
0	1	wednesday	5-3	G21	320[illegible]	LL	NOHY	B48	DE	64	25	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1	wednesday	4-2	G20	[illegible]340i xDrive	LL	NOHY	B58	US	71	34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
0	1	wednesday	4-2	G21	320d xDrive	RL	NOHY	B47	JP	51	31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1	1	tuesday	21-4	[illegible]	320i	RL	NOHY	B48	JP	45	28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

« Prev 1 2 3 4 5 6 7 8 9 10 Next »

Precision	Recall	Accuracy	F1
0.83	0.72	0.72	0.76

Classification Results (Confusion Matrix)

Predicted actual	Predicted 0	Predicted 1
0	[illegible] (73.6%)	[illegible] (26.4%)
1	[illegible] (37.1%)	[illegible] (62.9%)

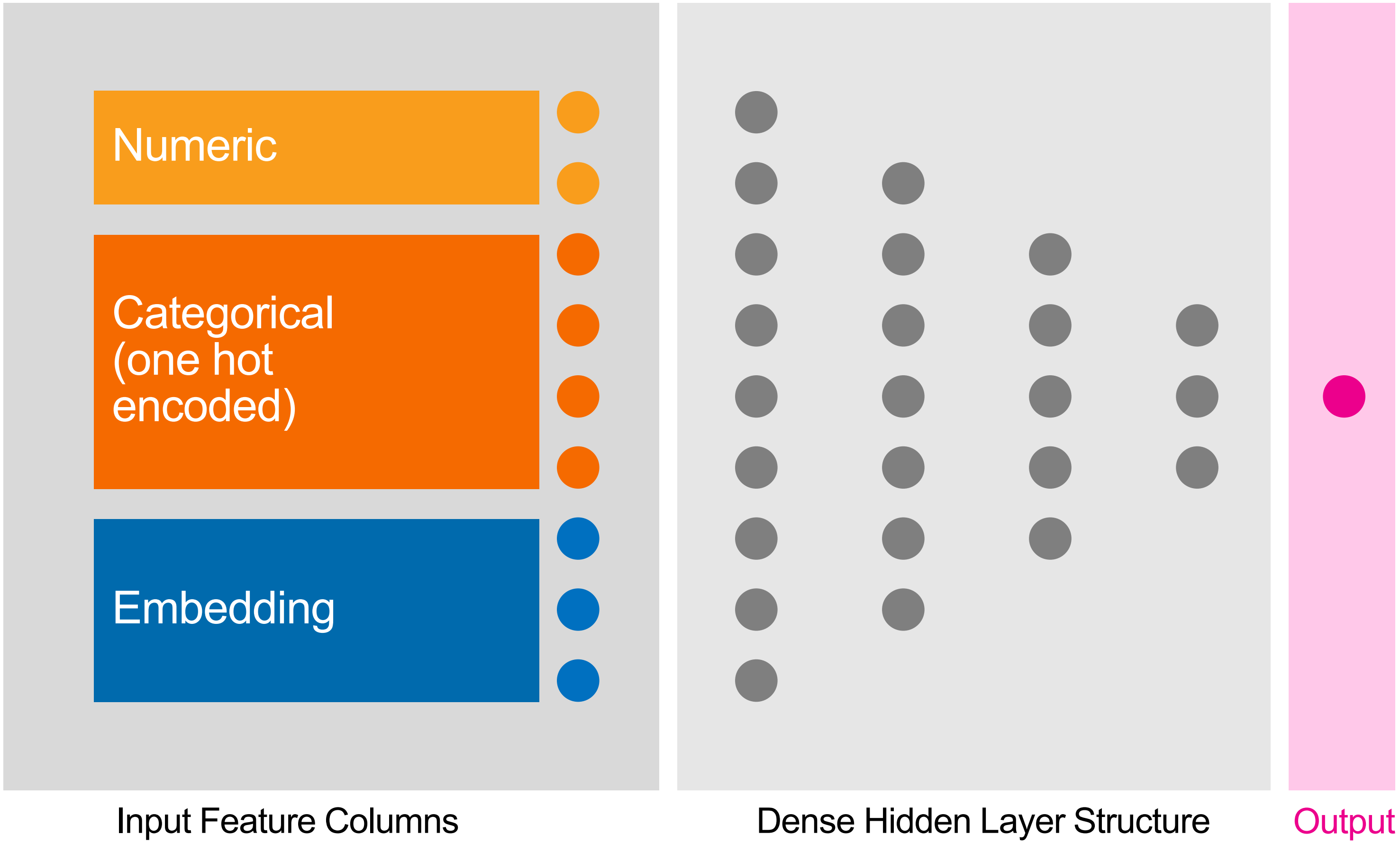
Note: All data used in the presentation are synthetically produced data.

Create Models

Improve with a Deep Learning Neural Network Approach

Deep Learning Approach

TensorFlow Neural Network Classifier with Feature Columns



More information: https://www.tensorflow.org/tutorials/structured_data/feature_columns

splunk> .conf20

Note: All data used in the presentation are synthetically produced data.

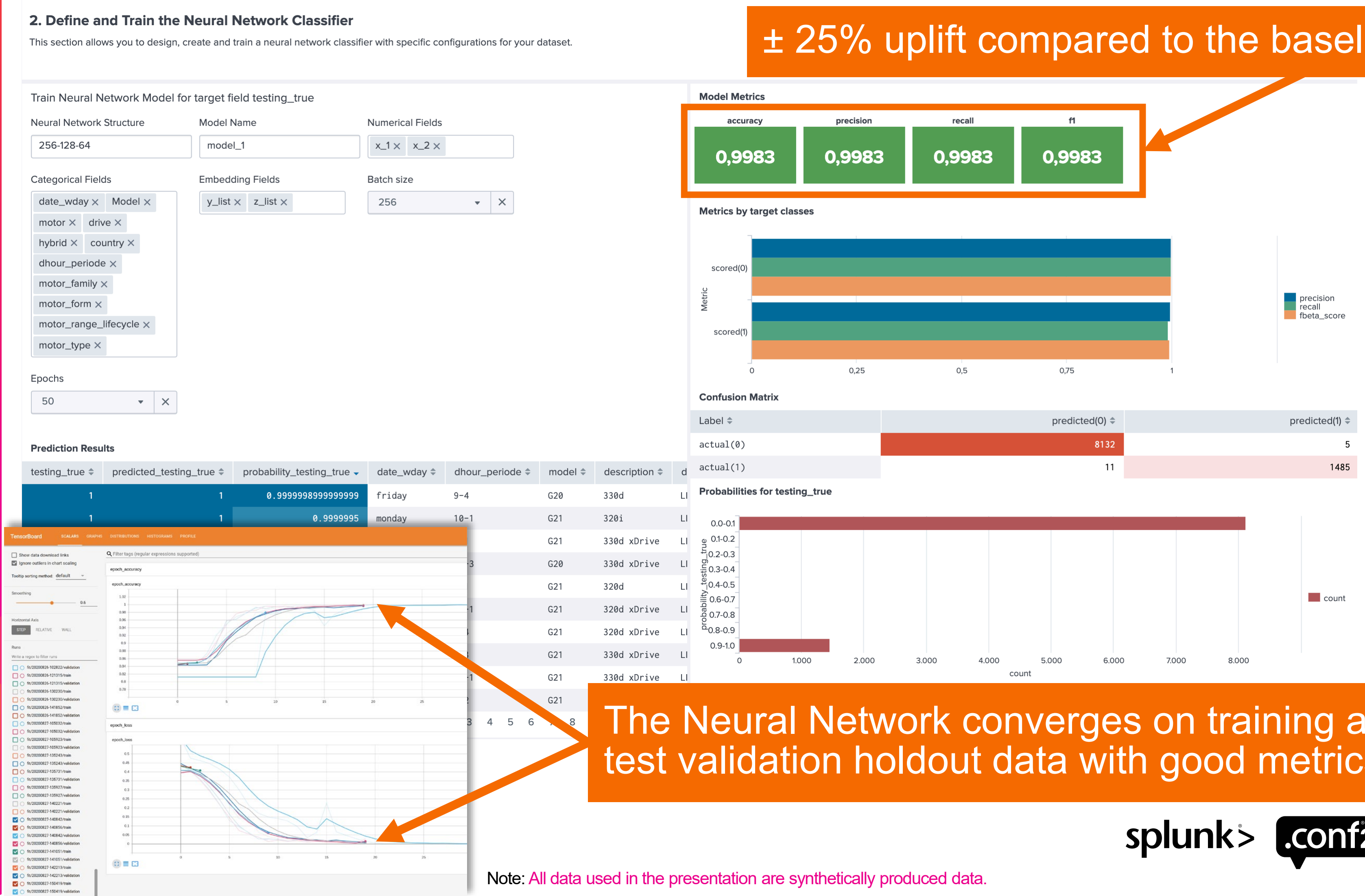
Create Models

Improve with a Neural Network and compare with existing baseline

Improvements with Deep Learning

Using a custom TensorFlow Neural Network Classifier

± 25% uplift compared to the baseline



The Neural Network converges on training and test validation holdout data with good metrics

Note: All data used in the presentation are synthetically produced data.

Apply Models

Apply the Neural Network and evaluate model performance on holdout data

Evaluate Model Performance

Using a custom TensorFlow Neural Network Classifier

Better than expected results and outperforming other existing approaches

Inference

Inference on Dataset

	testing_true	predicted_testing_true	probability_testing_true	date_wday	dhour_periode	model	description
1	0	1	0.99999833	tuesday	16-4	G20	330e
2	1	1	0.9999962	tuesday	21-4	G20	318d
3	0	1	0.99999523	monday	4-4	G21	M340i xDri
4	1	1	0.9999950999999999	monday	16-3	G21	320d
5	1	1	0.9999938	monday	11-3	G20	330d xDrive
6	1	1	0.99999166	tuesday	14-3	G21	330i xDrive
7	1	1	0.99998903	tuesday	8-3	G21	318d
8	1	1	0.9999886999999998	monday	10-2	G21	320d
9	1	1	0.99998844	thursday	18-3	G21	330d xDrive
10	0	1	0.9999882	tuesday	6-3	G21	330d xDrive
11	0	1	0.99998367	wednesday	14-4	G21	330d xDrive
12	0	1	0.99998343	monday	14-3	G20	330i
13	0	1	0.9999831	friday	12-4	G21	320d xDrive
14	0	1	0.99998224	friday	15-2	G21	320d
15	0	1	0.99998057	monday	11-3	G20	330d xDrive
16	0	1	0.99998045	tuesday	16-4	G20	330d xDrive
17	1	1	0.9999802	tuesday	16-4	G20	330e
18	0	1	0.9999795	wednesday	14-3	G21	330i
19	0	1	0.99997880[illegible]00001	monday	6-4	G20	320i
20	1	1	0.9999785[illegible]	monday	5-2	G20	330e
21	0	1	0.9999781999999999	mon[illegible]	6-3	G21	330i
22	1	1	0.99997795	friday	10-3	G20	320i
23	0	1	0.99997044	tuesday	18[illegible]	G21	330d xDrive
24	1	1	0.999969	friday	15-2	G21	320d
25	1	1	0.9999684999999999	tuesday	21-4	[illegible]21	330i xDrive

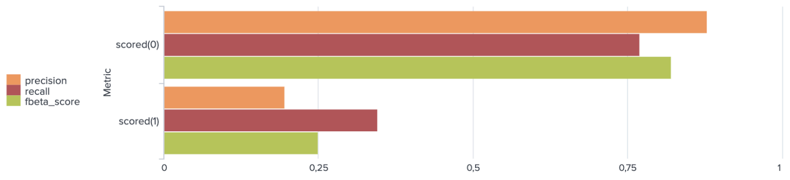
« Zurück 1 2 3 4 5 6 7 8 9 [illegible] Weiter »

Inference Scoring

Metriken

accuracy	precision	recall	f1
0,7109	0,7841	0,7109	0,7415

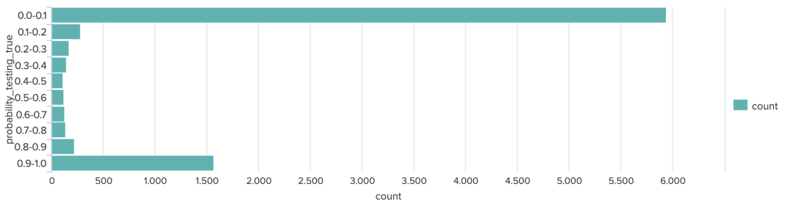
Metrics by target classes



Confusion Matrix

Label	predicted(0)	predicted(1)
actual(0)	5839	1745
actual(1)	802	425

Probabilities for testing_true



False positives still exist and further tuning and improvement is possible

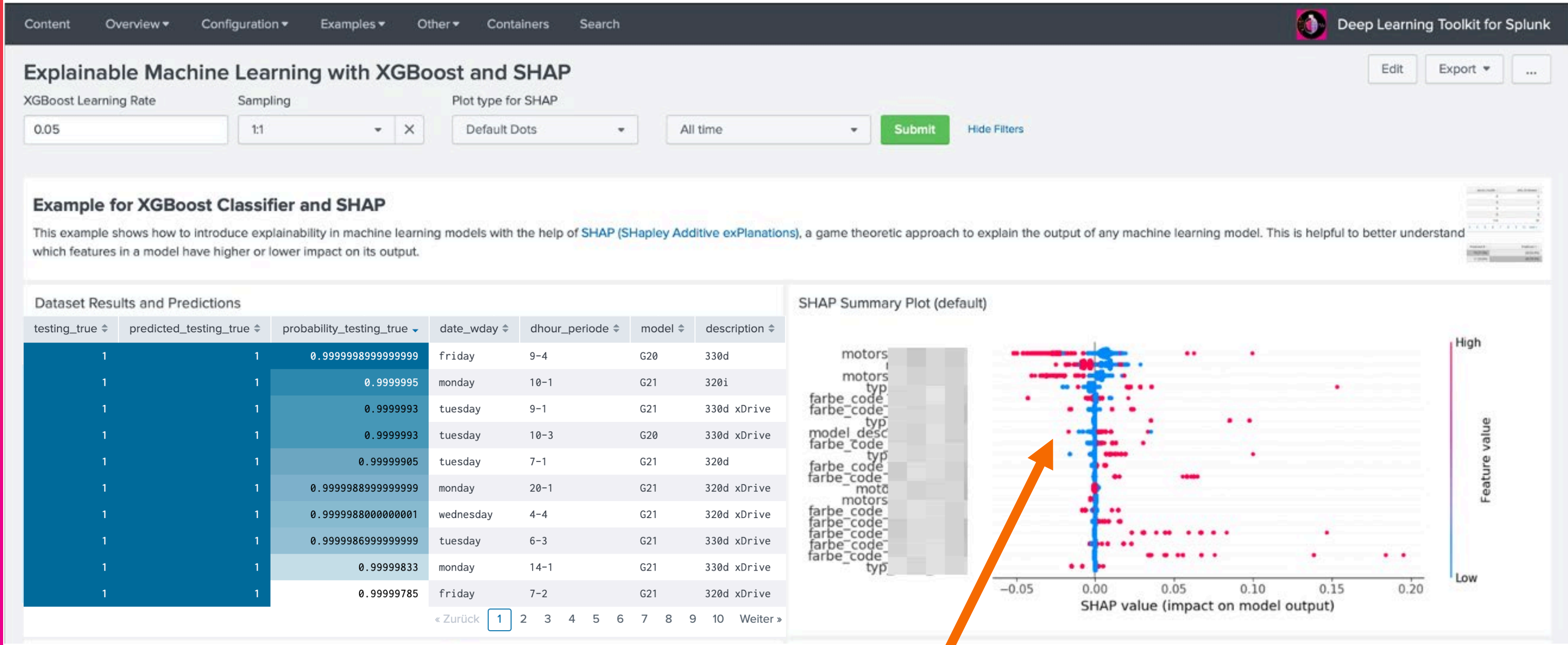
Note: All data used in the presentation are synthetically produced data.

Evaluate and Understand

Better understand how a machine learning model is working and which features have what impact on a model's output

Model Explainability for Better Transparency and Additional Insights

Make Machine Learning Models Interpretable



testing_true	predicted_testing_true	probability_testing_true	date_wday	dhour_periode	model	description
1	1	0.9999998999999999	friday	9-4	G20	330d
1	1	0.9999995	monday	10-1	G21	320i
1	1	0.9999993	tuesday	9-1	G21	330d xDrive
1	1	0.9999993	tuesday	10-3	G20	330d xDrive
1	1	0.99999905	tuesday	7-1	G21	320d
1	1	0.9999988999999999	monday	20-1	G21	320d xDrive
1	1	0.9999988000000001	wednesday	4-4	G21	320d xDrive
1	1	0.9999986999999999	tuesday	6-3	G21	330d xDrive
1	1	0.99999833	monday	14-1	G21	330d xDrive
1	1	0.99999785	friday	7-2	G21	320d xDrive

SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model.

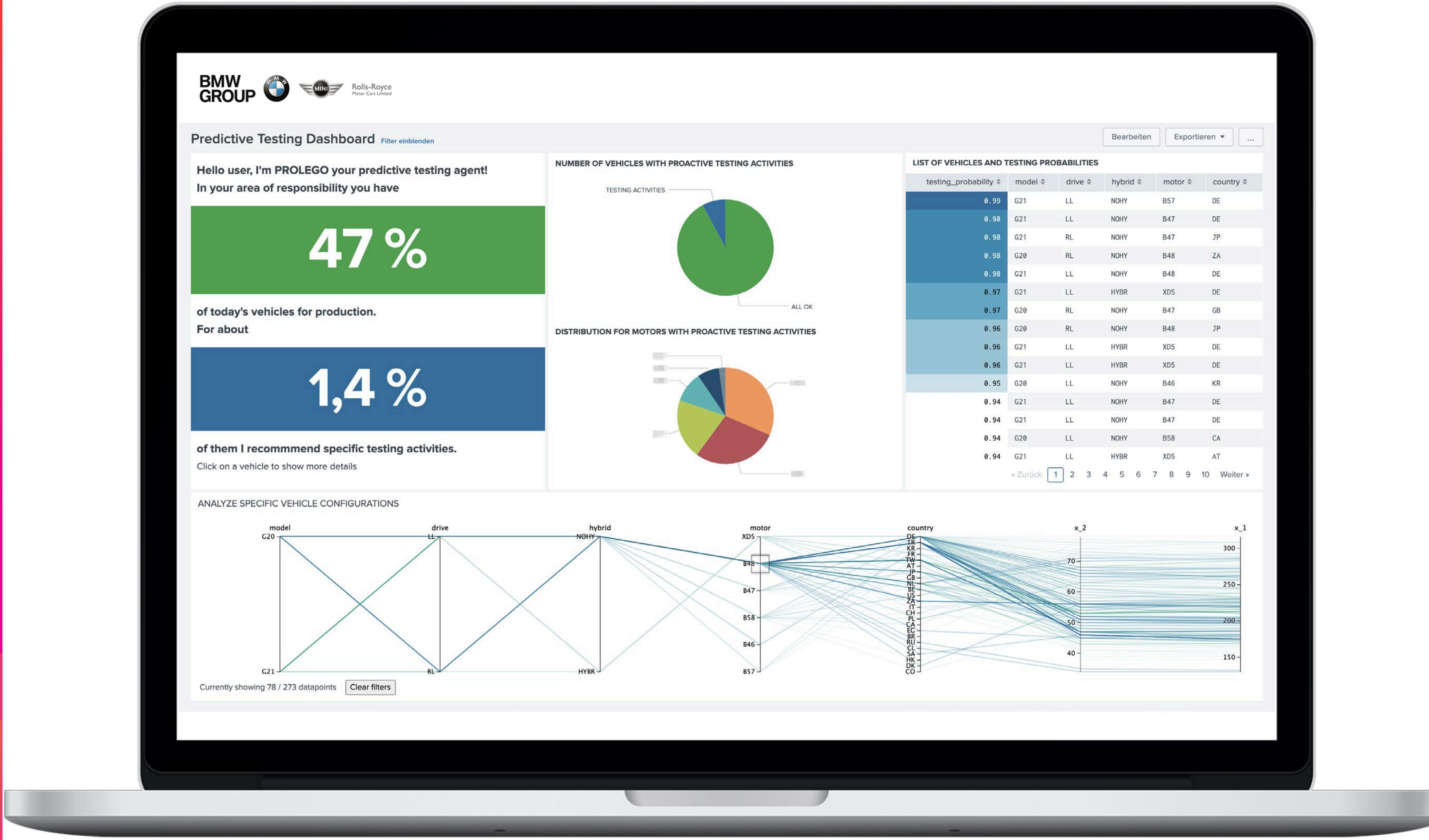
More information on SHAP: https://github.com/slundberg/shap

Note: All data used in the presentation are synthetically produced data.

Present and Operationalize

Make the model results accessible and usable in a simples and easy way

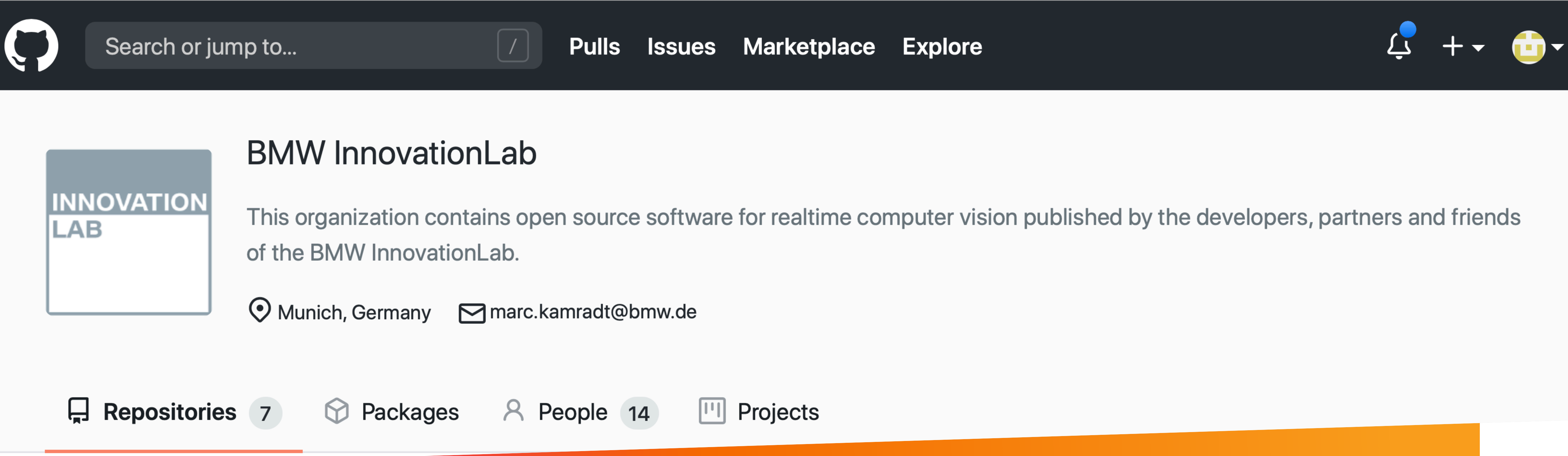
Final Dashboard to Communicate Results with Drill Downs for Details



Note: All data used in the presentation are synthetically produced data.

Future

What’s Next?



Sharing is Caring

Learn more about BMW Innovation Lab

With Splunk and DLTK: Sky is the Limit!

Thank You

Please provide feedback via the

SESSION SURVEY