Yogurt Sale Prediction

Customer Data Analytics – Migros

Group 6 – Case 1

29.10.2021

Natalie Broter

Jonathan Hahn

Loris Lindenmann

Adrian Peng

Sinah Schüpfer

Beatrice Stierli

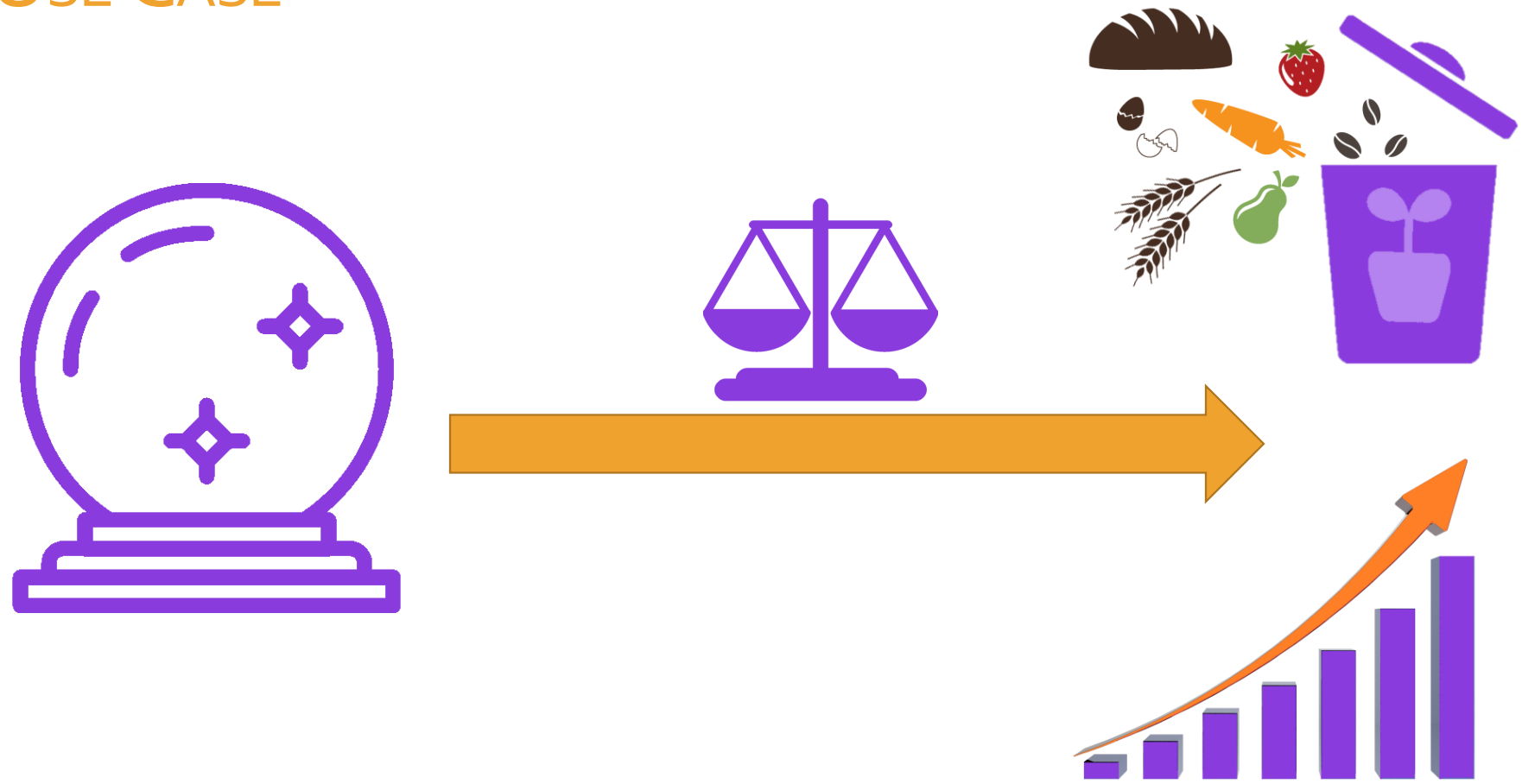
Raphael Suter

SEVEN-UP DATA FACTORY

Agenda

- Research Question - Goal
- Data Basis & Data Enrichment
- Models for Sale Prediction
- Results
- Reflection & Dealing with Problems
- Implications

Use Case



Research Question

How can we find the best model to predict sales data for multiple products on the example of yogurts?

DATA

WHICH FACTORS INFLUENCE SALES?

"Healthier lifestyle, more knowledge about food and higher income is associated with more yoghurt intake"

(D'Addezia et al., 2015; Possa et al, 2020; Robinson 2017)

SALES

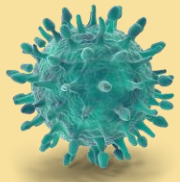
PRICES

Google TRENDS

"During lockdown on average one kilogramm more yoghurt in Switzerland"

(BauernZeitung Donnerstag, 25. März 2021)

EVENTS



Data as Basic for our Solution

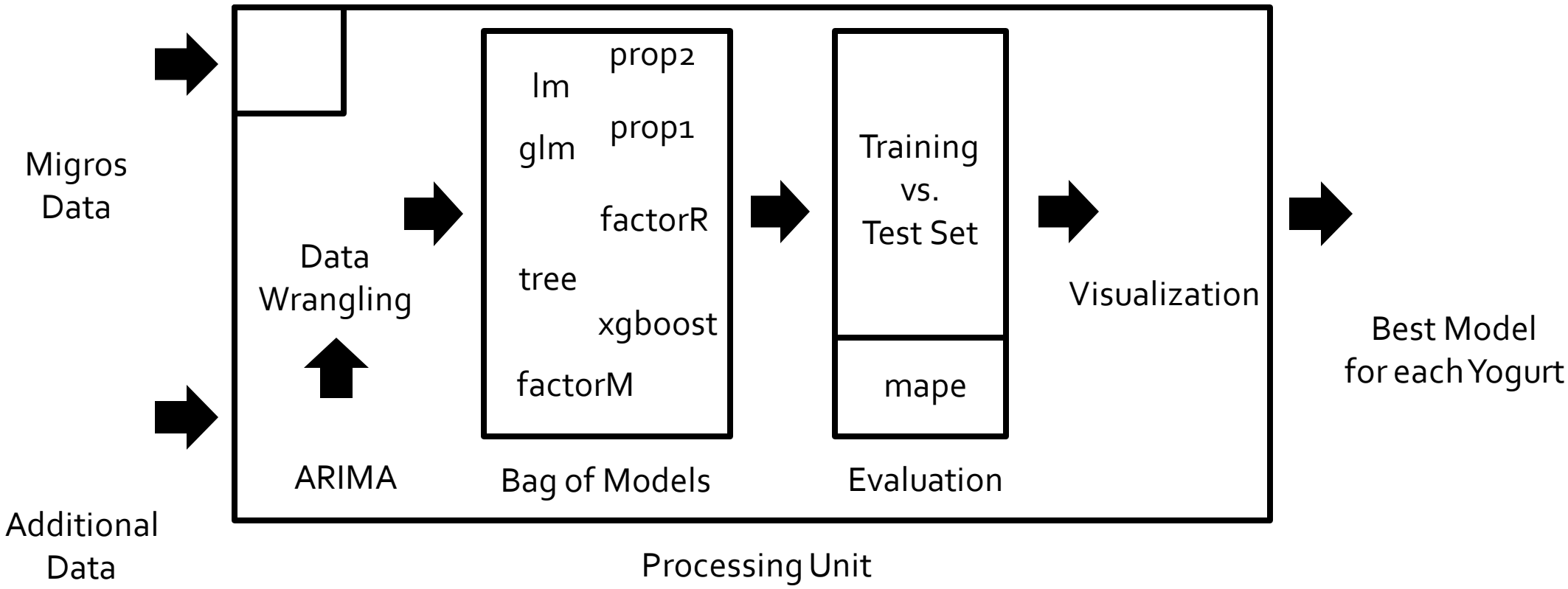
Data Migros	Consolidate Data for Models	Enrich Data
Article		Yogurt prices
Year / month / week		Milk prices
Sales		Google search terms „gesund.essen“ & „joghurt“
Promotion 1-5		Lockdown dates
		Holidays

PROCEDURE & ARCHITEKTURE

Problem

SOLUTION

Procedure – Architecture



MODELS & RESULTS

Best Model for each Yogurt

lm	glm	tree	prop1	prop2	xgboost	factorM	factorR
0	0	1	1	2	4	10	1

prod	model
BIO JOGHURT NATURE 180G	RegressionTree
YOGOS GRECQUE NATURE 180G	Prophet1
BIO FAIRT.JOG.MOKKA 180G	Prophet2
EXC JOGHURT TRUFFES 150G	Prophet2
M-CLAS JOG. NATURE 200G	XGBOOST
M-CLAS JOG. VANILLE 200G	XGBOOST
BIFIDUS JOGH. NATURE 500G	XGBOOST
VALFLORA CREME FRAICHE NA	XGBOOST
MB JOGHURT NATUR 500G	ProphetScoresM
M-CLAS JOG. HEIDELBE 200G	ProphetScoresM

prod	model
M-CLAS JOG. HIMBEER 200G	ProphetScoresM
M-CLAS JOG. SCHOKOLA 200G	ProphetScoresM
M-CLAS JOGHURT MOKKA 200G	ProphetScoresM
M-CLAS JOG. APF/MANG 200G	ProphetScoresM
M-CLAS JOG. ERDBEER 200G	ProphetScoresM
M-CLAS JOG. HASELNUS 200G	ProphetScoresM
BIO JOGHURT NATURE 500G	ProphetScoresM
BIFIDUS JOGH. NATURE 150G	ProphetScoresM
AHA JOGHURT LAKTOSEF CLAS	ProphetScoresRest

Zoom in Excellence Truffes – MAPE in Testset

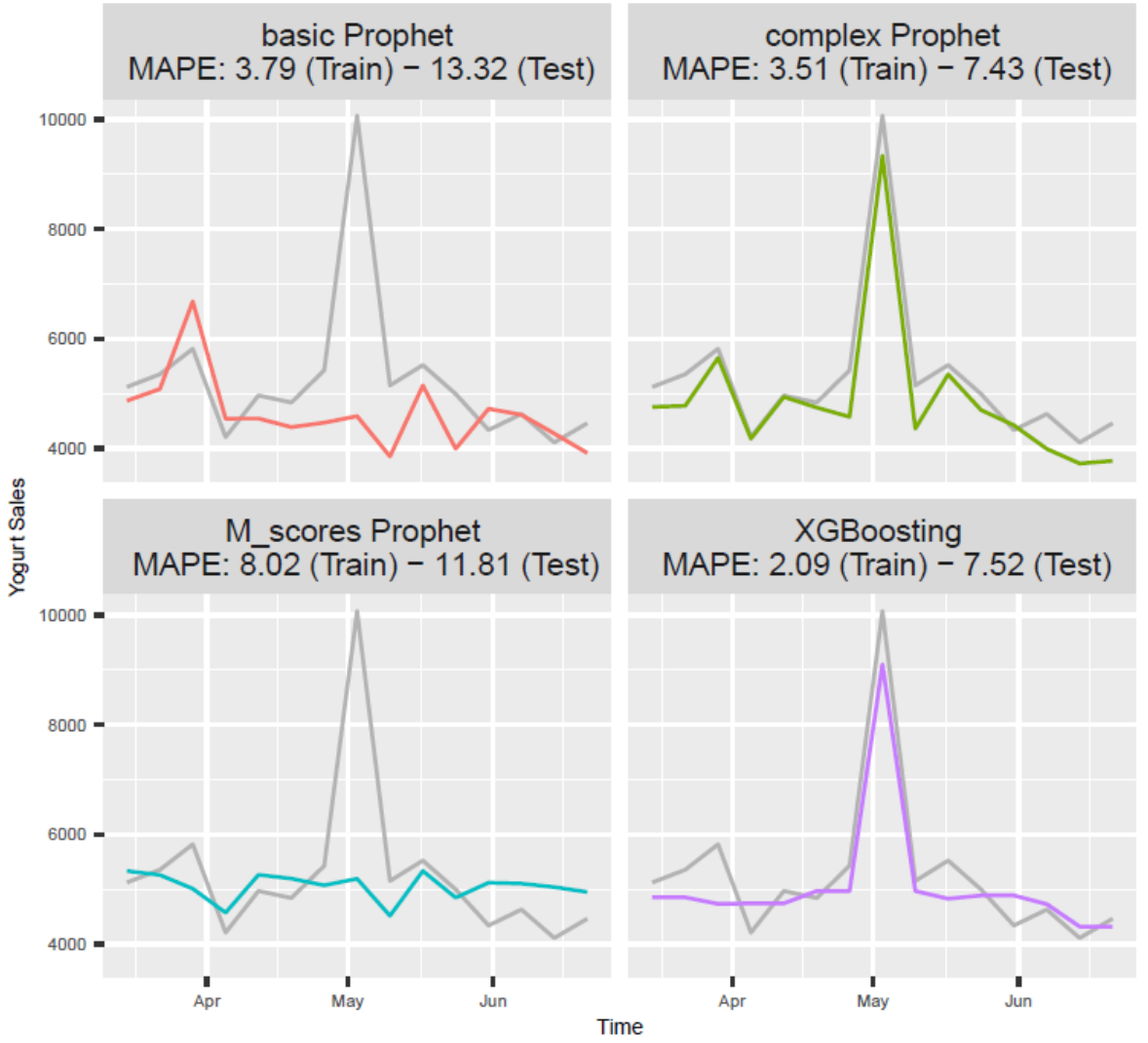
Excellence Truffes 150g

	lm	glm	tree	prop1	prop2	xgboost	factorM	factorR
BIO FAIRT.JOG.MOKKA 180G	12.16	11.32	11.59	8.32	5.42	7.08	11.06	10.37
MB JOGHURT NATUR 500G	11.84	15.01	7.79	8.76	8.21	6.52	5.95	9.56
M-CLAS JOG. HEIDELBE 200G	9.05	12.46	11.55	11.23	9.32	7.32	5.51	8.53
M-CLAS JOG. NATURE 200G	9.54	11.76	10.59	8.05	7.30	5.38	13.95	15.62
M-CLAS JOG. HIMBEER 200G	9.24	14.30	13.17	17.71	14.54	7.30	6.46	11.22
M-CLAS JOG. SCHOKOLA 200G	11.19	9.86	7.16	8.59	10.62	8.72	6.20	6.50
M-CLAS JOGHURT MOKKA 200G	11.00	10.23	7.27	8.43	9.23	10.82	5.89	7.54
M-CLAS JOG. APF/MANG 200G	10.62	11.88	7.86	9.90	9.00	8.38	6.52	7.18
M-CLAS JOG. ERDBEER 200G	9.06	13.76	5.64	10.92	8.29	8.86	3.55	7.38
M-CLAS JOG. VANILLE 200G	13.60	14.68	9.60	11.02	10.59	7.60	8.21	9.64
M-CLAS JOG. HASELNUS 200G	9.16	10.37	10.14	9.28	8.49	9.73	7.75	10.64
BIFIDUS JOGH. NATURE 500G	7.05	9.71	5.24	8.37	5.71	4.85	16.45	6.86
BIO JOGHURT NATURE 180G	9.64	9.82	6.51	11.36	11.34	7.82	6.63	14.42
EXC JOGHURT TRUFFES 150G	15.30	13.74	20.03	13.32	7.43	7.52	11.81	11.31
YOGOS GRECQUE NATURE 180G	7.30	11.70	13.14	5.16	6.06	6.72	20.42	9.07
BIO JOGHURT NATURE 500G	12.70	17.77	7.49	13.32	11.53	9.06	7.24	15.35
BIFIDUS JOGH. NATURE 150G	10.25	10.04	6.13	9.41	7.08	5.91	5.82	8.53
VALFLORA CREME FRAICHE NA	16.50	21.34	13.71	21.06	14.04	9.22	18.08	11.71

lm	glm	tree	prop1	prop2	xgboost	factorM	factorR
15.30	13.74	20.03	13.32	**7.43**	**7.52**	11.81	11.31

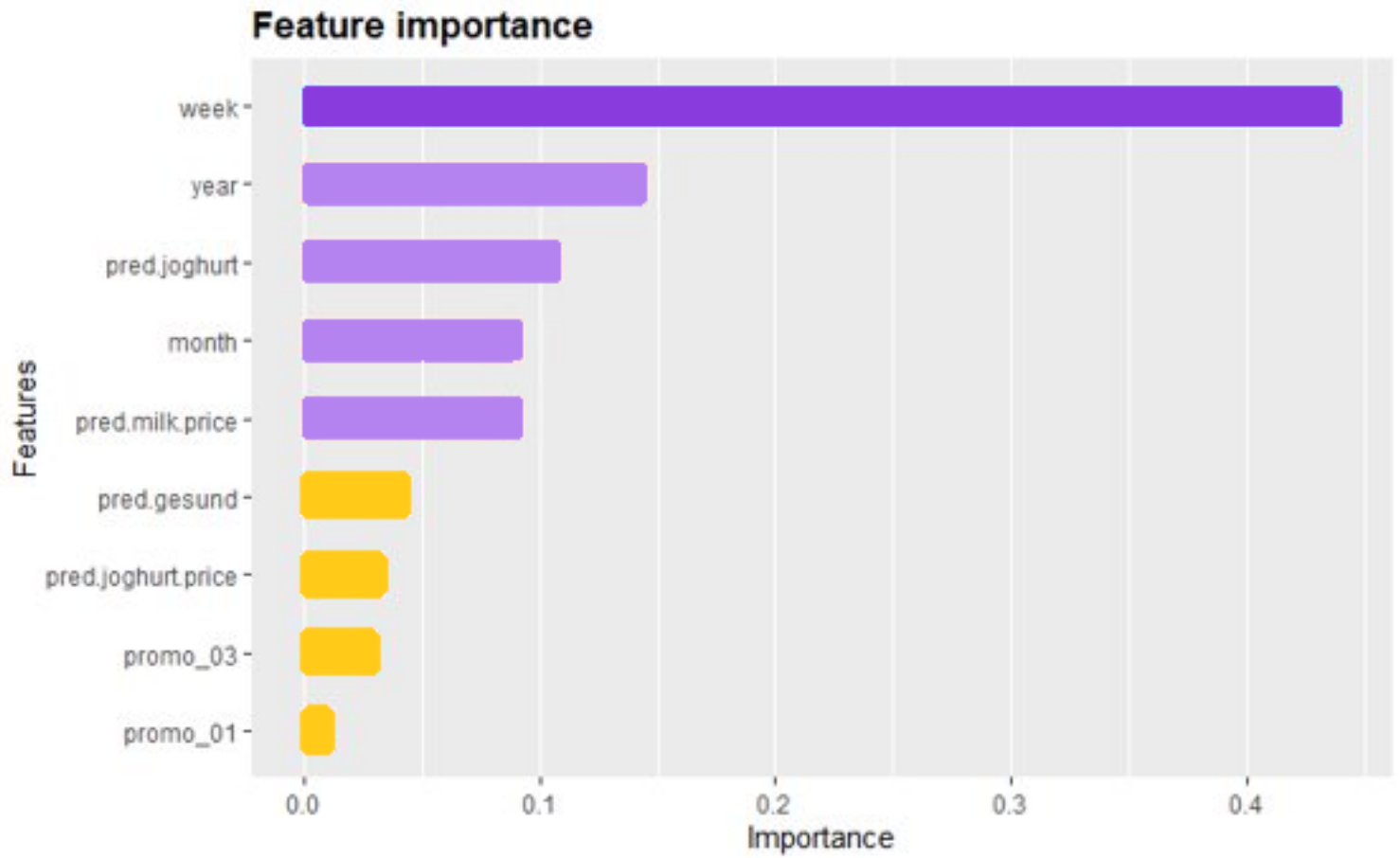
Zoom in Exellence Truffes – Model Comparison

Excellence Truffes 150g



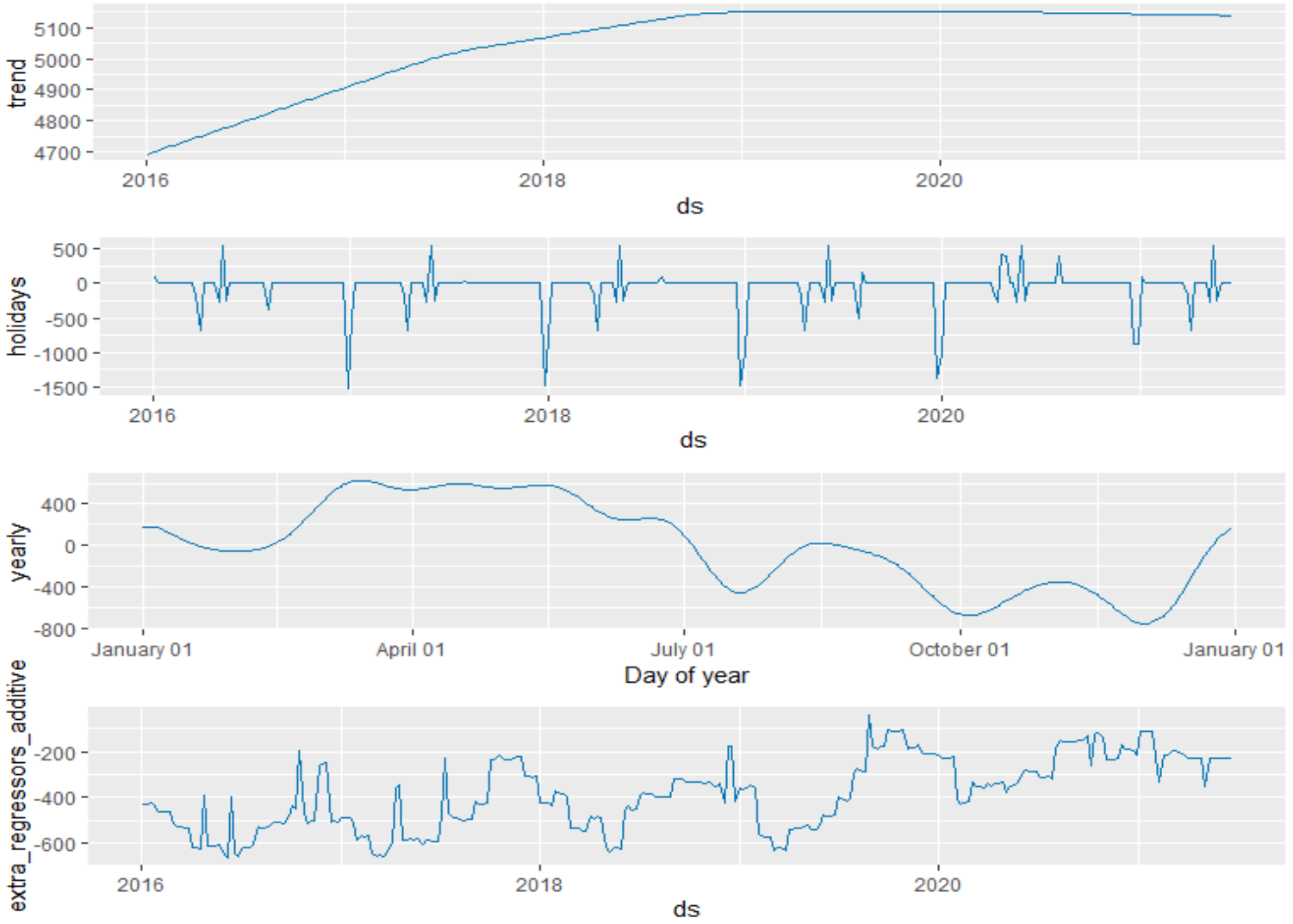
ZOOM IN EXCELLENCE TRUFFES – XGBOOSTING

Excellence Truffes 150g



Zoom in Excellence Truffes- Prophet2

Excellence Truffes 150g



Additional Features

- Automatically predict future values with best model (based on MAPE of Testset)
- Other models or multiple models > specify in script
- PDF with code and all important plots

REFLECTION & IMPLICATIONS

Reflection – Data

GTrends API

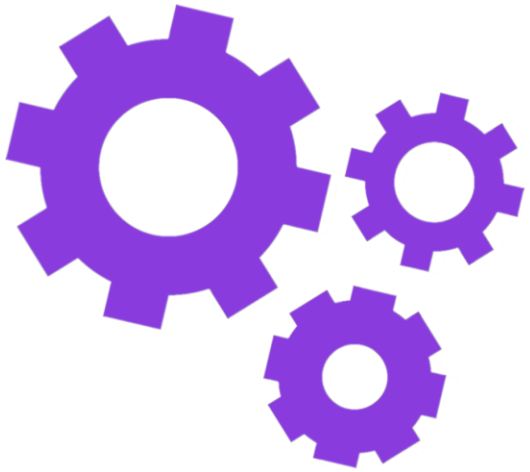
- API retrieves time interval dependent on input
- Limited amount of API calls per day -> Future: get data in a separate micro service

Yogurt & Milk Prices

- just monthly
- no yogurt prices from Migros

Other data like Cumulus or data per store

Reflection – Procedure & Models



Regressors must be predicted

- Google Trends, Milk & Yogurt Prices
- Bad predictions can strongly worsen the result

Model Evaluation

- Cross validation only for one time interval
- Better: cross validation on rolling basis

Any model for each yogurt

- Not always the same method worked best
- LM/GML models not optimized for the individual yogurt

IMPLICATIONS

SALE
FORECAST

added value:
automatically calculate all models for each yogurt and predicts with the best one

adapted production ->
reduction of
production costs & food waste

Outlook

OPTIMIZE ?

Evaluation of Forecast

Cost Benefit Analysis

QUESTIONS

Sources

- https://www.bauernzeitung.ch/artikel/landleben/pro-kopf-wurde-2020-rund-ein-kilo-mehr-joghurt-verspeist-353195
- D'Addezio, L., Mistura, L., Sette, S., & Turrini, A. (2015). Sociodemographic and lifestyle characteristics of yogurt consumers in Italy: Results from the INRAN-SCAI 2005-06 survey. Mediterranean Journal of Nutrition and Metabolism, 8(2), 119-129.
- Possa, G., de Castro, M. A., Marchioni, D. M. L., Fisberg, R. M., & Fisberg, M. (2015). Probability and amounts of yogurt intake are differently affected by sociodemographic, economic, and lifestyle factors in adults and the elderly—results from a population-based study. Nutrition Research, 35(8), 700-706.
- Robinson, R. A. (2017). Examination of the Demand for Yogurt by Brand, 2009-2011 (Doctoral dissertation).