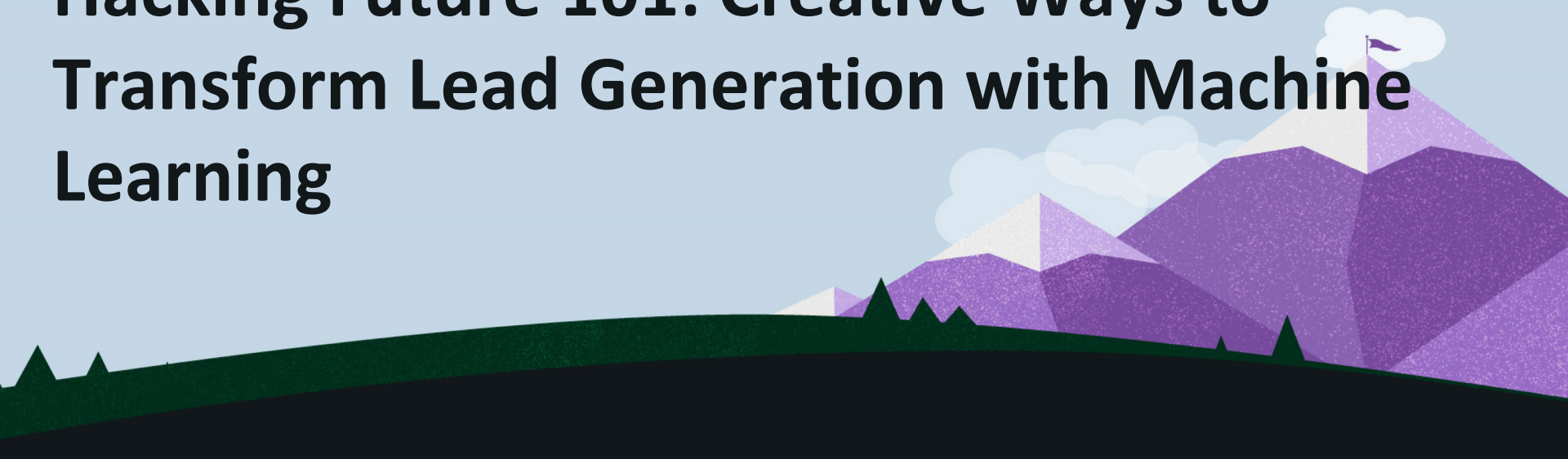


SESSION:

# Hacking Future 101: Creative Ways to Transform Lead Generation with Machine Learning

A stylized, low-poly landscape illustration in shades of purple and green. The foreground is a dark green silhouette of a hill with several small, dark green triangular shapes representing trees. The background features several purple mountains of varying heights and colors, some with a lighter purple or white peak. A small purple flag is visible on the highest peak. The sky is a light blue gradient.

# Hacking Future 101: Creative Ways to Transform Lead Generation with Machine Learning

Advanced | Tech



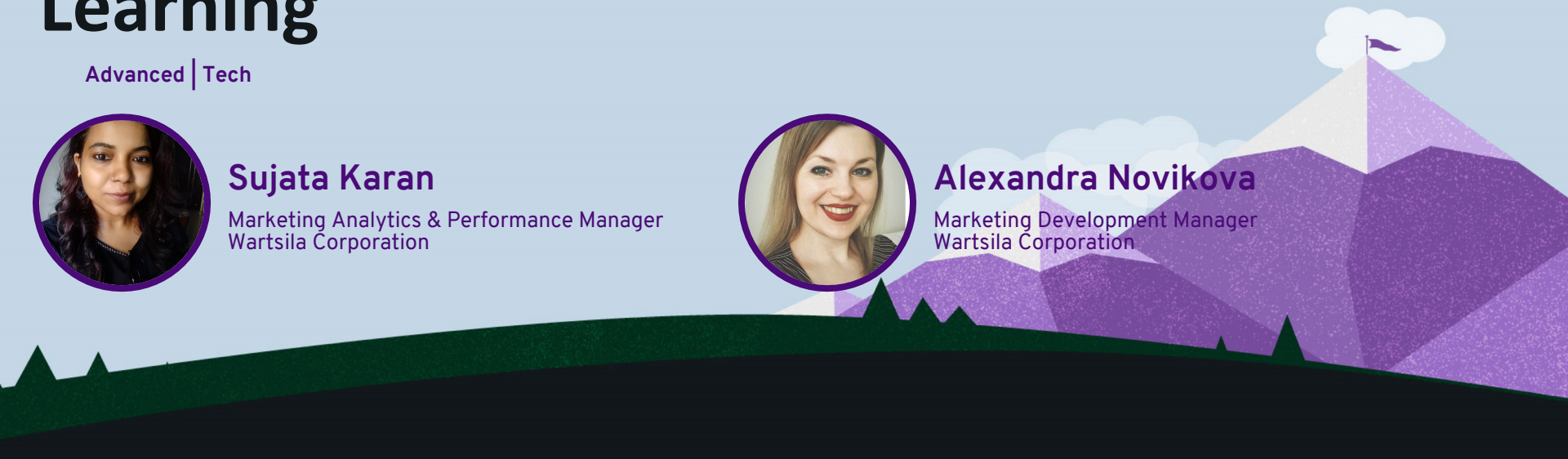
**Sujata Karan**

Marketing Analytics & Performance Manager  
Wartsila Corporation



**Alexandra Novikova**

Marketing Development Manager  
Wartsila Corporation

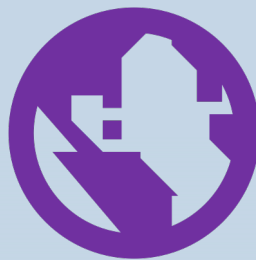


# Who We Are

WÄRTSILÄ Corporation



4.6 Mn Euro



70 Countries

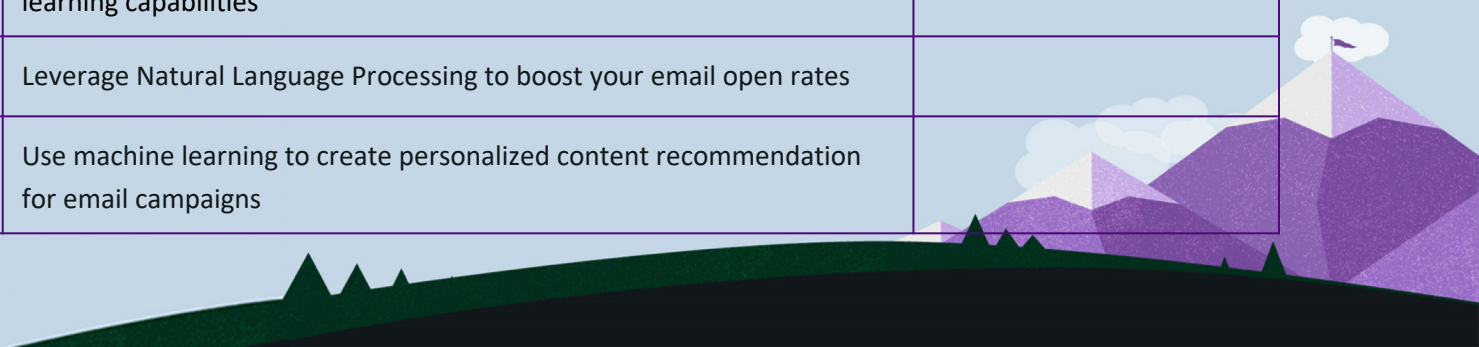


18000 Employees  
140 Nationalities

# Speaker Agenda

What's going on here?

Time	Agenda	Extra Info
10:30-10:45	Use data and machine learning to increase the impact of your marketing program	
10:45-11:00	Segment your customers in conditions of scattered or missing data	
11:00-11:15	Improve customer targeting and campaign response with machine learning capabilities	
11:15-11:20	Leverage Natural Language Processing to boost your email open rates	
11:20-11:25	Use machine learning to create personalized content recommendation for email campaigns	

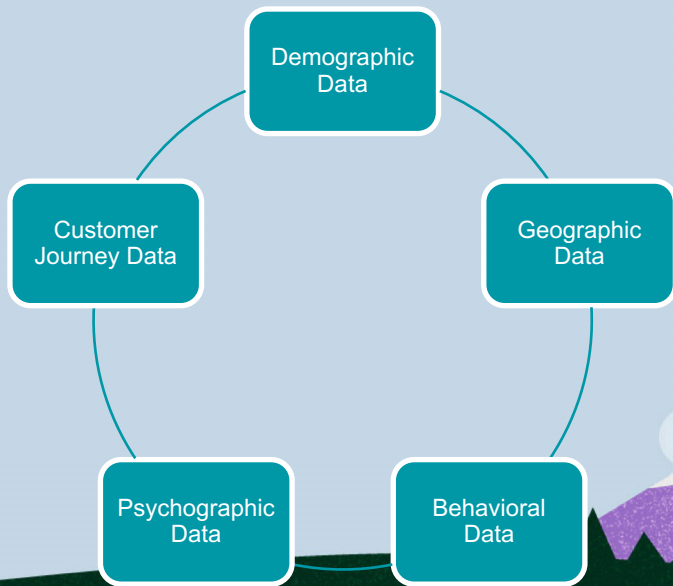


# Marketing that can be improved with ML

- Recommendation System
- Target Audience Prediction
- Lifetime Value Prediction
- Personalization with NLP
- SEO Automation & Optimization



# Scattered or Missing data?



# Scattered or Missing data?

## Audit and taking stock of data

- Correct
- Clean
- Complete
- Properly Formatted (standardized)
- Verified

## Remove clutter from data

- Duplicate Data
- Irrelevant data
- Redundant Data
- Inaccurate Data
- Low Quality Data

## Know & Evaluate Your CRM Data Collection Methods

- Customer-Facing Data Input Forms
- Employee-Facing Data Input Forms
- Third-Party Customer List Imports
- Integrations

## Define Your CRM Data Standards

## Standardize Data

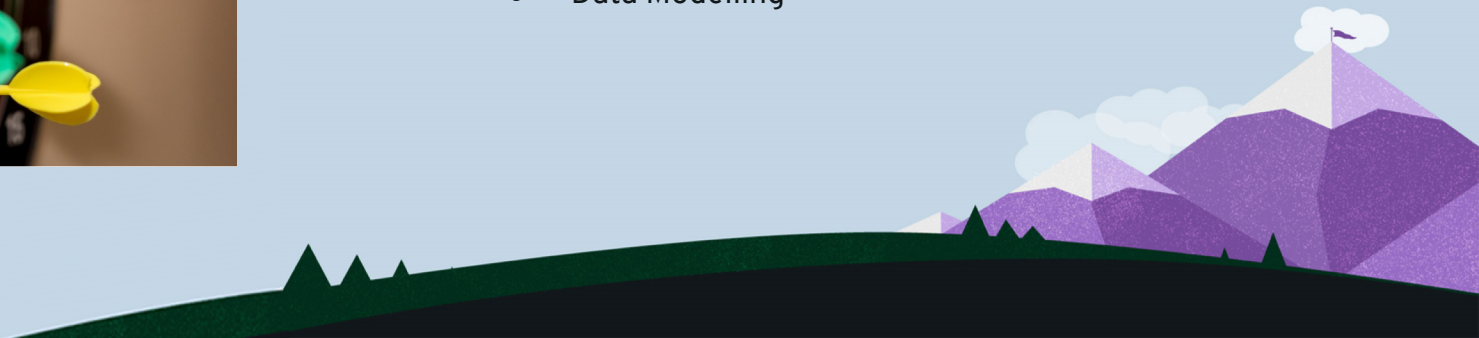


# Improve targeting and campaign response with ML learning capabilities

Probability ~60%  
Accuracy ~ 75%



- Exploratory Data Analysis
- Outlier Detection
- Feature Engineering
- Feature Selection
- Data Modelling





# Exploratory Data Analysis

- Missing data fields
- Discrimination ability of each field
- Correlation of the discrete variable
- Distribution according to campaign response

Campaign ID	Contact Id	HASRESPO NDED	Account Record	Account Status	Line Of Business	Market Segment	Role	account_ record_ty	contact_ departm	role-2	title
1337	3962	1	ART-1	AS-1	LOB-1	MS-1	Owner	Customer	Managem	Decision M	Admistrad
1596	2480	1	ART-2	AS-2	LOB-2	MS-2	Owner	Customer			Admistrad
1935	3828	1	ART-3	AS-3	LOB-3	MS-3	Trader	Customer	Managem	Decision M	Commodit
1716	3371	1	ART-4	AS-4	LOB-4	MS-4	Trader	Prospect			Commodit
1529	3373	1	ART-5	AS-5	LOB-5	MS-5	Engineering co.	Customer	Other	Influencer	Environm
1757	3900	1	ART-6	AS-6	LOB-1	MS-6	Agent	Customer	Managem	Decision M	General M
1292	2017	1	ART-7	AS-7	LOB-2	MS-7	Owner	Customer			General M
1607	2018	1	ART-8	AS-1	LOB-3	MS-8	Agent	Customer			Superinte
1039	2519	1	ART-9	AS-2	LOB-4	MS-9	Management	Customer	Material	Supply Unit	Material S
1891	3467	1	ART-10	AS-3	LOB-5	MS-10	Management	Customer	Material	Supply Unit	Material S
1032	2093	1	ART-1	AS-4	LOB-1	MS-11	Other	Customer			Material S
1968	3936	1	ART-2	AS-5	LOB-2	MS-12	Other	Customer	Managem	Decision M	Sub Secret
1445	2313	1	ART-3	AS-6	LOB-3	MS-13	Management	Supplier			
1905	3859	1	ART-4	AS-7	LOB-4	MS-14	Owner	Customer		Decision Maker	

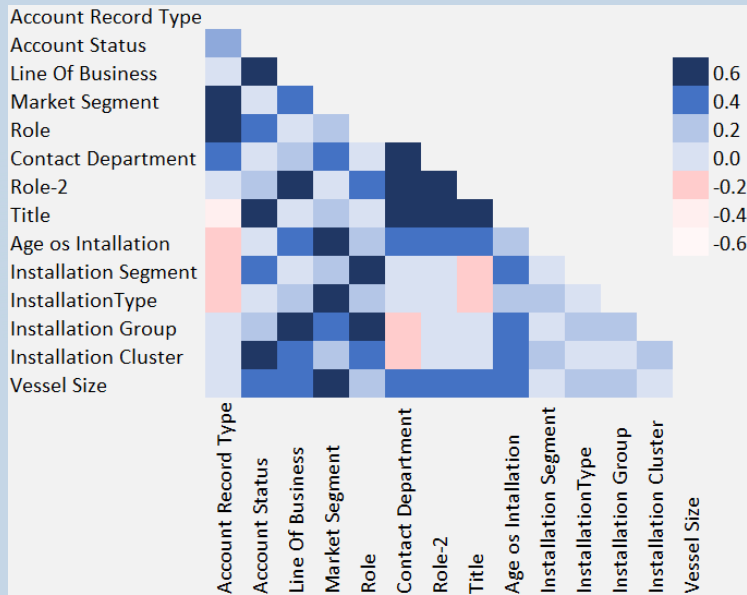
# Exploratory Data Analysis Cont'd

- The discrimination ability stands for how well the category correlates to the target
- Number of data points
- Arbitrability setting the discrimination ability to 15% and 40 observation

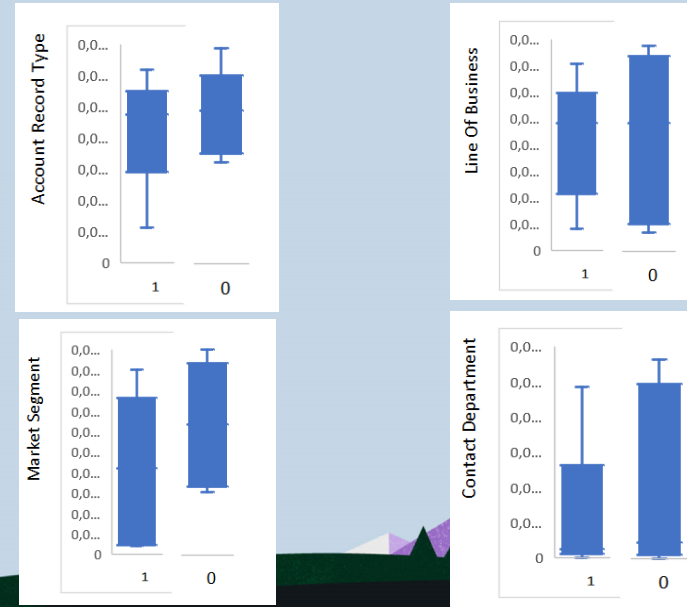
		Number of Observation	Discrimination Ability	More than n Observations
Account Record Type	ART-1	153	<b>0,108374</b>	TRUE
	ART-2	896	0,16765	TRUE
	ART-3	64	<b>0,03768</b>	TRUE
	ART-4	48	0,189083	TRUE
	ART-5	12	<b>0,018768</b>	<b>FALSE</b>
Line of Business	LOB-1	480	0,206897	TRUE
	LOB-2	567	0,246546	TRUE
	LOB-3	198	0,176538	TRUE
	LOB-4	10	0,5	<b>FALSE</b>
	LOB-5	2	0,568711	<b>FALSE</b>
Market Segment	MS-1	280	0,208374	TRUE
	MS-2	172	0,26765	TRUE
	MS-3	915	<b>0,13768</b>	TRUE
	MS-4	83	0,289083	TRUE
	MS-5	67	<b>0,118768</b>	TRUE
	MS-6	31	0,306897	TRUE
	MS-7	499	0,346546	TRUE
	MS-8	586	0,276538	TRUE
	MS-9	217	<b>0,086781</b>	TRUE
	MS-10	29	0,668711	<b>FALSE</b>
	MS-11	21	0,308374	<b>FALSE</b>
	MS-12	299	0,36765	TRUE

# Exploratory Data Analysis Cont'd

Correlation matrix of the discrete variables

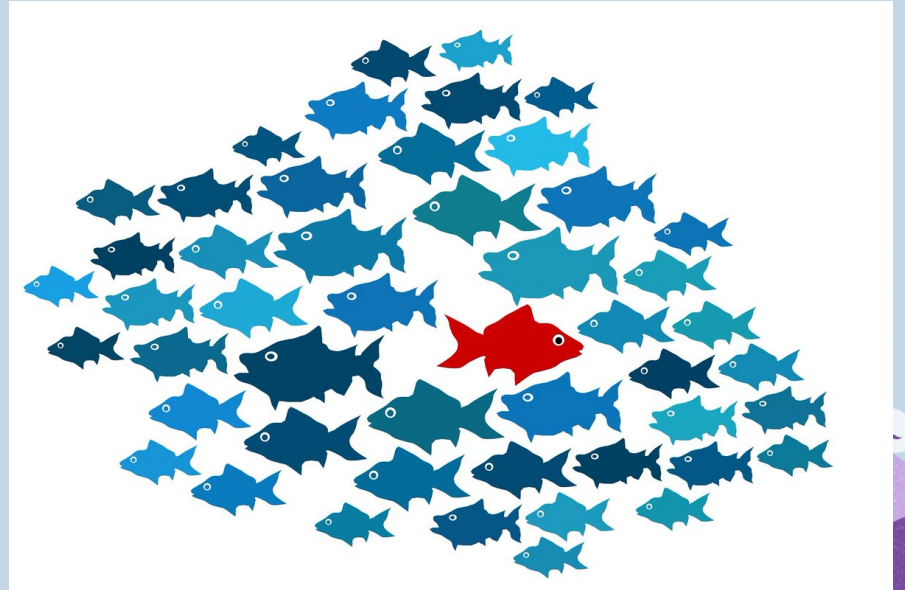


Value distribution according to campaign response



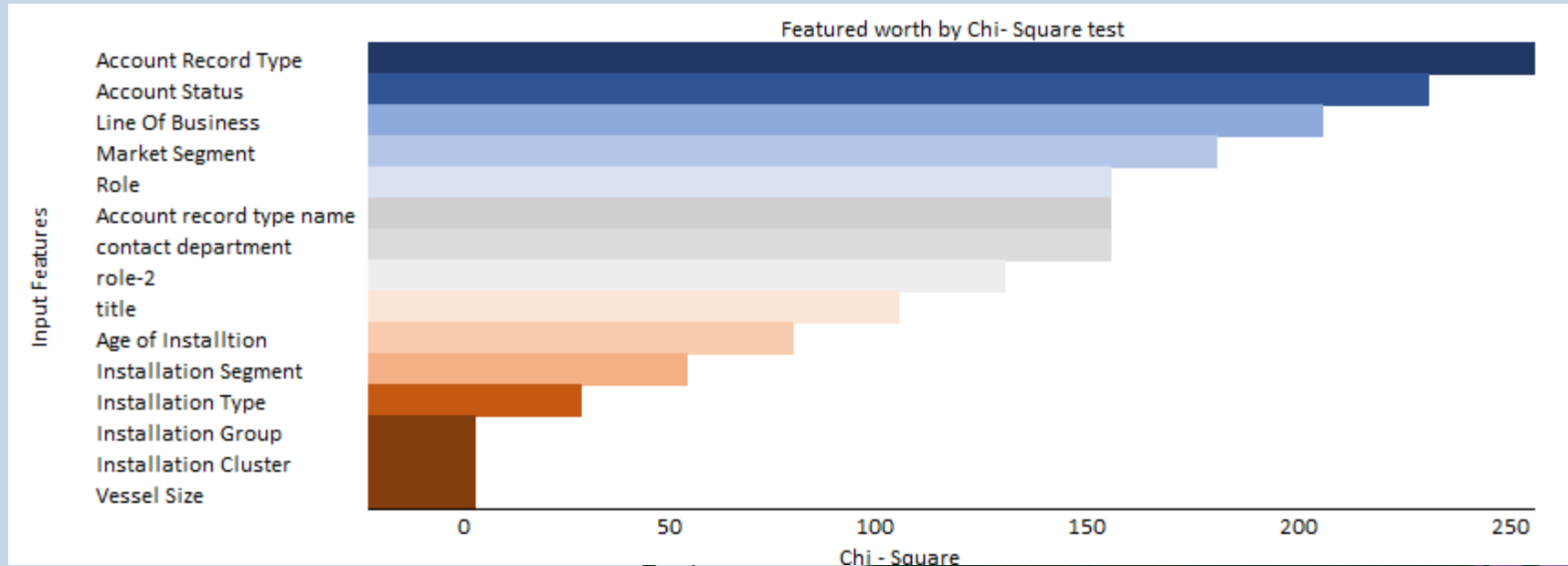
# Outlier Detection

- Data entry errors
- Measurement errors
- Experimental errors
- Intentional
- Data processing errors
- Sampling errors
- Natural



# Feature Engineering

Feature selection is the process of selecting a subset of relevant features for use in model construction

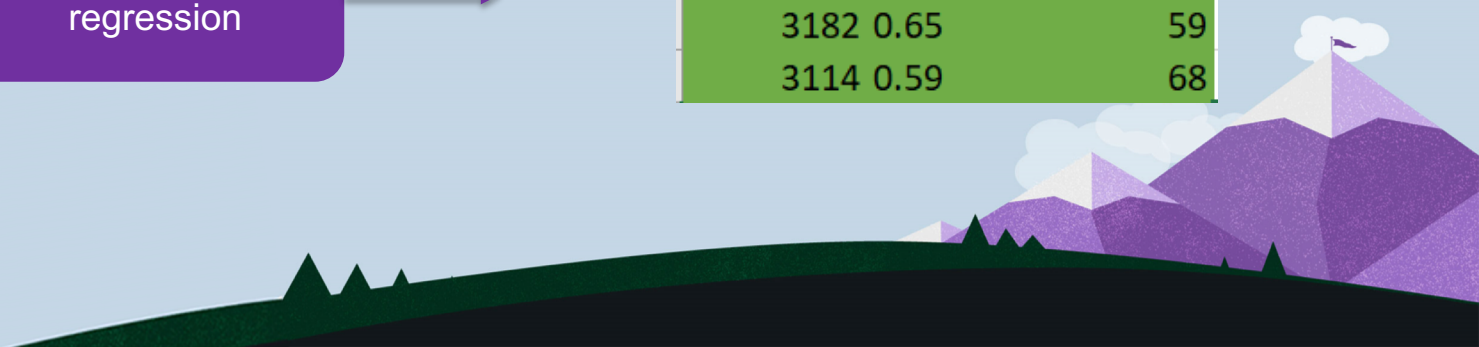


# Data Modelling

Oversampling

Logistics regression

Contact Id	Probability	Accuracy
3604	0.30	68
2626	0.67	65
2809	0.85	64
2517	0.44	44
2138	0.20	40
2783	0.18	53
2706	0.37	47
3182	0.65	59
3114	0.59	68



# What is NLP?

- NLP helps computers to understand human language with ML and deep learning techniques
- It takes an unstructured data and turns it into a structured format by utilizing:
  - Entity recognition
  - Sentiment analysis
  - Topic modelling
  - Text classification
  - Etc.
- Subtopics of NLP:
  - NLU (machine reading comprehension via grammar and context)
  - NLG (text generation based on provided data)

## Entity Recognition

In 1961, Walt Disney Productions licensed certain film and other rights of Milne's Winnie-the-Pooh stories from the estate of A. A. Milne and the licensing agent Stephen Slesinger, Inc., and adapted the Pooh stories, using the unhyphenated name "Winnie the Pooh", into a series of features that would eventually become one of its most successful franchises.

### Tags:

NAME DATE CHARACTER

## Sentiment Analysis

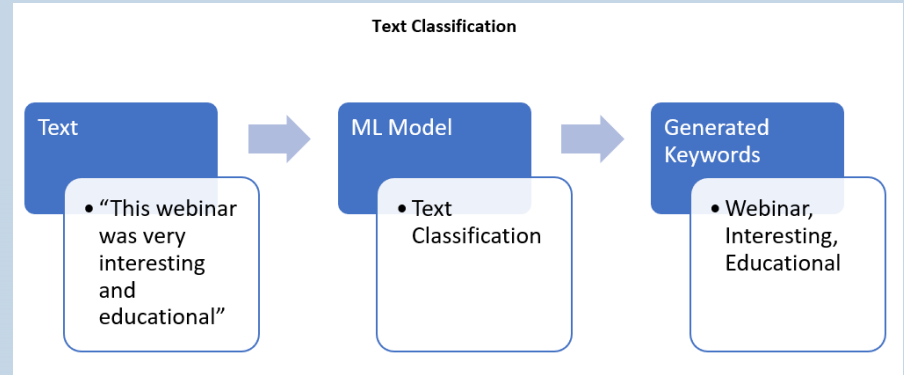
"I didn't like the webinar" – *Is a negative sentiment*

"I didn't hate the webinar" – *Is a neutral sentiment*

"I liked the webinar" – *Is a positive sentiment*

# Common Use Cases for NLP

- Content creation
  - NLG can create keyword-optimized blog posts, wepages, emails, reports etc.
- Sentiment analysis
  - Comments/reviews/competitors data
- SEO optimization
  - Keywords and topics discovery, content analysis and optimization/creation, tagging images etc.
- Customer segmentation
  - Based on jargon and keywords
- Chatbots
  - Customer interaction
- Email classification
  - Spam/ad detection





# ML and NLP role in improving Email Campaigns rates

- Usage of NLP techniques can be utilized to improve Open Rate and CTR of email campaigns:
  - Detecting most commonly used keywords/topics/sentiments/customer intent in the highly opened emails' subjects and email content
  - Discovering potential target groups/topics based on email content
  - Grouping the findings and providing recommendations for content creators to create highly personalized email content
- As a second step, once data will be reliable enough, NLP can be used in auto-generating possible content for email campaigns



# Example: Email Content Analysis

**Number**   **Benefit**   **Topic**

To: michael.jensen@company.com

Subject: 4 Creative Ways How to Improve Your Email Marketing Efficiency with Machine Learning

Hi Michael,

Did you know that 65% of consumers do not open emails that start with letter A?

In this [article](#) "Utilize Machine Learning to Boost your Email Open Rate" we explore how to utilize various machine learning models to improve your email campaign metrics. In this 10 min read, we'll cover:

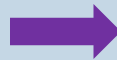
- Boosting engagement
- Customer segmentation with Machine Learning
- Text analysis with Machine Learning
- Sending right emails to right customers

Click here to start reading it:

**ACCESS NOW**

Hope you enjoy it!

Kind regards,  
Anne Green



**Prospect Group:** Marketing Manager

**Topics:** Machine Learning, Marketing, KPI

**Type of campaign:** Webinar invitation, Whitepaper

**Email sentiment:** Informal

**Subject line:** Max 20 characters

**Subject line sentiment:** Informal

**Subject line characters:** Question mark, Exclamation mark

**Subject line keywords:** marketing efficiency, improve marketing KPI, marketing management

**Email content:** Informal

**Email length:** 100

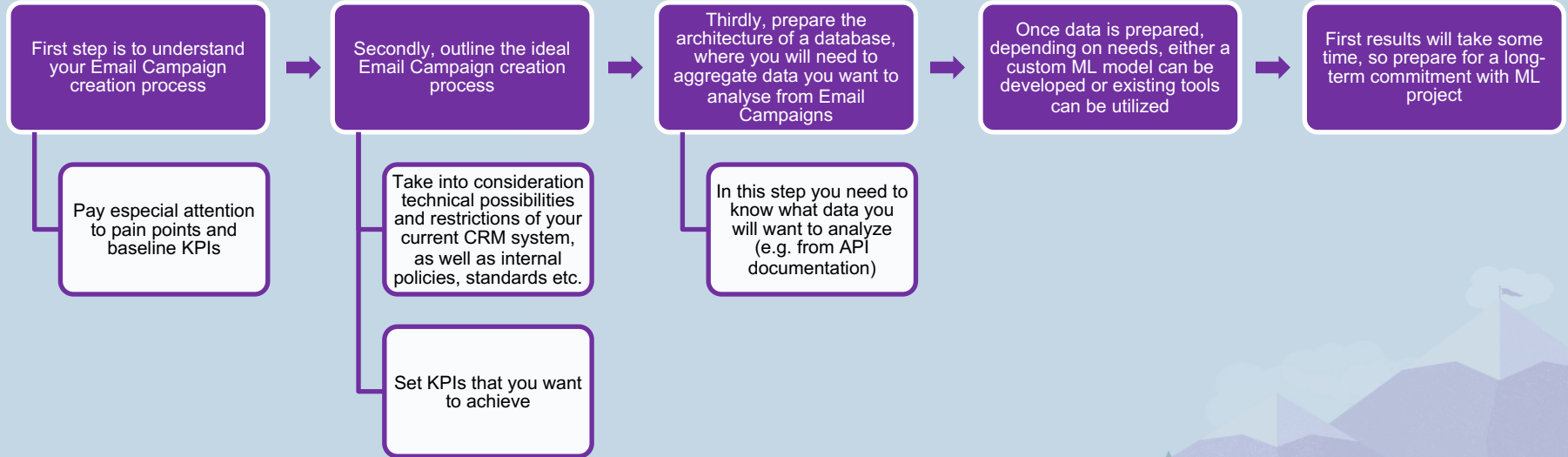
**Email content keywords:** marketing automation, marketing strategy, online advertising, types of marketing, market research

**#of CTA:** 1

**Media:** Video, Photo

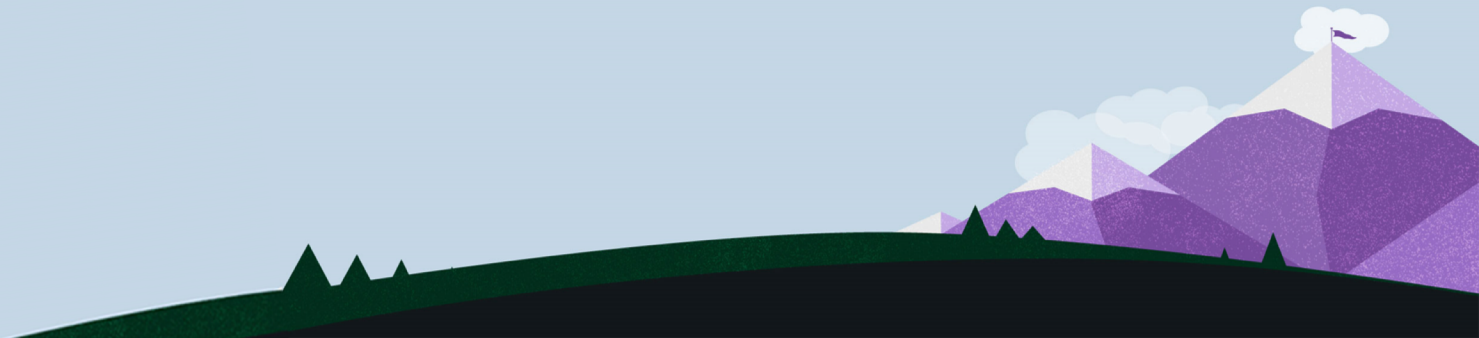
**Bullet point/plain text:** Bullet point

# How to Make it Happen?



# KPIs to Measure the model success

- Open rate
- Click to Open ratio
- Click-through rate
- Increased number of participants in webinar



# Thanks for listening!



Sujata Karan    Alexandra Novikova

You can find us on LinkedIn or hit us up at  
[sujata.karan@wartsila.com](mailto:sujata.karan@wartsila.com) and  
[alexandra.novikova@wartsila.com](mailto:alexandra.novikova@wartsila.com)





# Thank You To Our Sponsors

#PARDREAMIN2021 |  
@PARDREAMIN

## TITLE SPONSORS



# SERCANTE

## PLATINUM SPONSORS



feedotter

SAASCEND



terminus



GetFeedback™



QUALIFIED

## GOLD SPONSORS

velir

THE SPOT  
FOR PARDOT by SERCANTE

## SILVER SPONSORS

FIONTA

CLOUDTEGIC