

Disrupting Lead Generation through Artificial Intelligence

Bisnode Group Analytics

Pierre Deville – Head of Data Science and Analytics

Artificial Intelligence

“

Machines that can solve a problem
which requires human intelligence

”





data

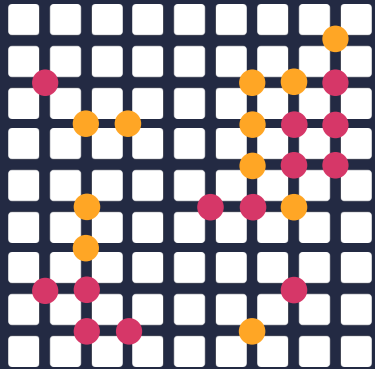
Artificial Intelligence

“

Machines that can solve a problem
which requires human intelligence

Solved by perceiving contextual information

”



What about Lead Generation ?

Usual approach



Official firmographic
data



Structured data



Static information



Low granularity



Traditional Statistics



Need for **accurate** and **dynamic** company
pictures

How to leverage contextual information to offer game changing performances ?

Key Aspects



Focus on **unstructured online data**
Where valuable information actually is



Connect **AI paradigms** together
Crucial to combine variety of unstructured data



Automation & Scalability
Across geographies but also languages



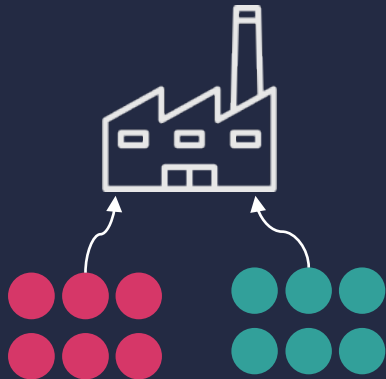
Machine Learning Pipeline

Overview

ENRICH

LEARN

PREDICT



94%

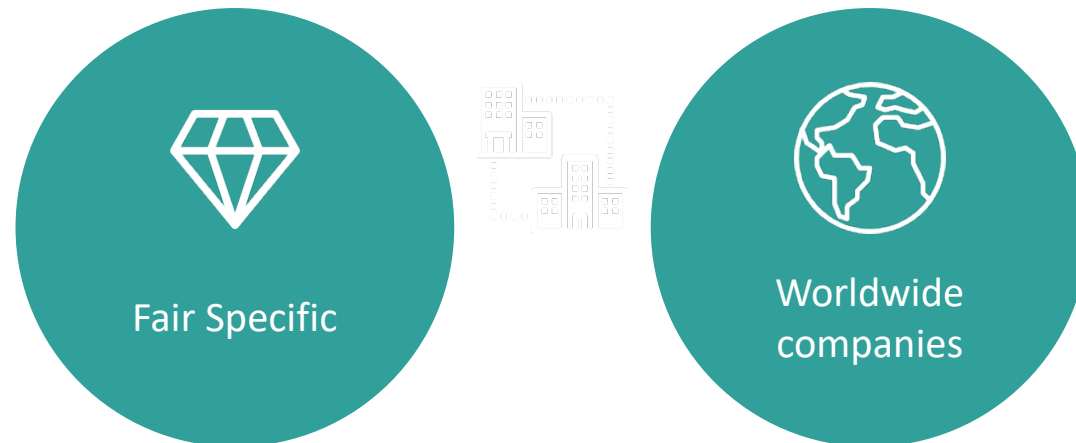
Use case



World's largest trade fair, congress
and event organizer

Objective

Detect high quality leads



techtextil

14 - 17.05.2019

Frankfurt am Main, Germany

CONNECTING THE FUTURE



Agrotech



Buildtech



Clothtech



Geotech



Homitech



Indutech



Medtech



Mobiltech



Oekotech



Packtech



Protech



Sporttech



in parallel with:

texprocess



messe frankfurt

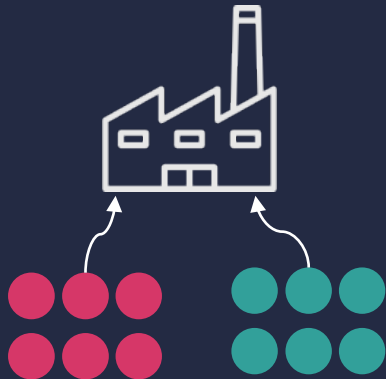
Machine Learning Pipeline

Overview

ENRICH

LEARN

PREDICT



94%

Machine Learning Pipeline

Overview

ENRICH Website Content

Company linkages

Products

Partnerships

Global Reach

Publications

Images

Jobposts

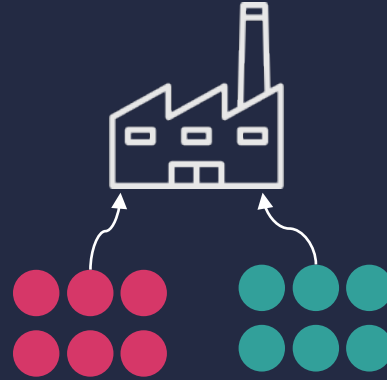
Company online information can be identified through:

- Website identification
 - Gradient Boosting on website features derived search engine results
- Company mentions
 - Named Entity Recognition combined with Elastic Search disambiguation on 1M+ news articles

Machine Learning Pipeline

Overview

ENRICH



Website Content ●
Company linkages ●
Products ●
Partnerships ●
Global Reach ●
Publications ●
Images ●
Jobposts ●


11M+ Companies


39 Countries

● Company name
● Industry codes
● Company age
● Company size
● Revenue
● Profit/loss
● Equity
● Subsidiaries

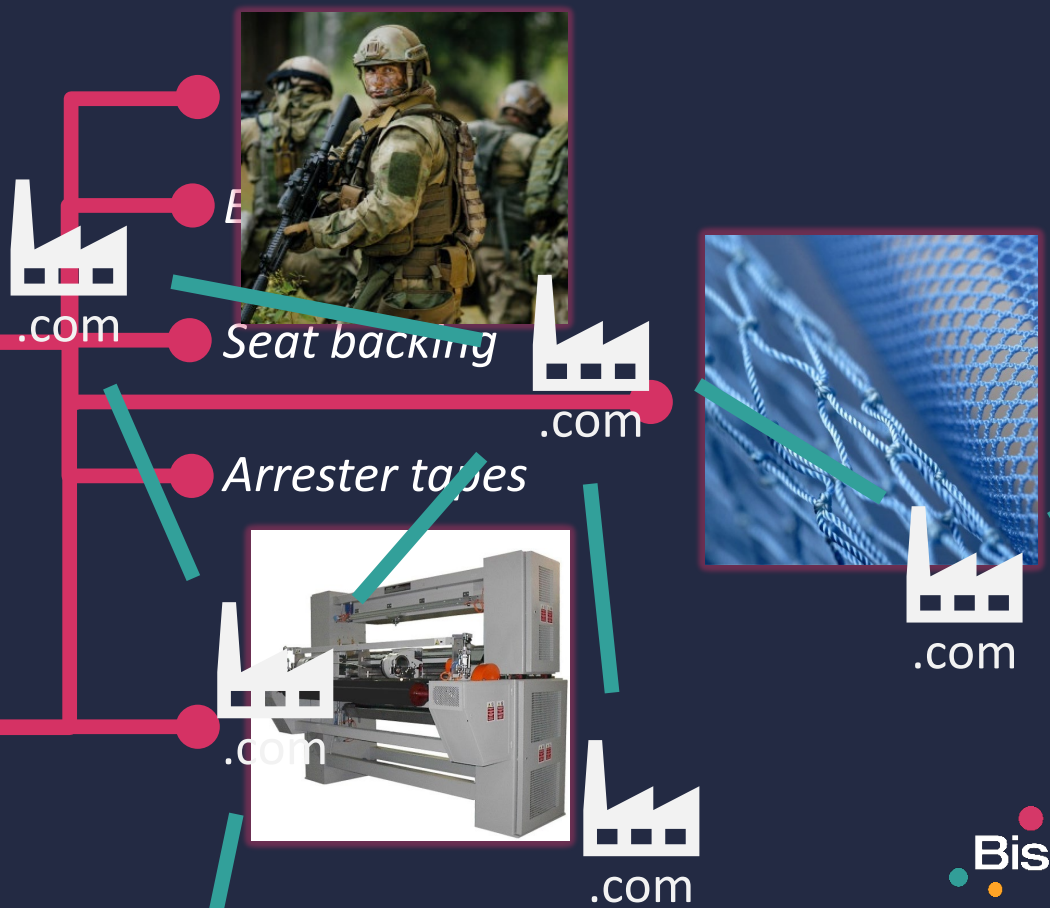
Machine Learning Pipeline

ENRICH

1. Automated feature extraction from unstructured data

LEARN

- Website Content
- Company linkages
- Products
- Partnerships
- Global Reach
- Publications
- Images
- Jobposts



Core identifying to identify company for Image recognition

Not a 2D Bayes Classifier to derive features from the network

Machine Learning Pipeline

LEARN

1. Automated feature extraction from unstructured data

Website Content ●
Company linkages ●
Products ●
Partnerships ●
Global Reach ●
Publications ●
Images ●
Jobposts ●

Naïve Bayes Classifier to
derive respective
probabilities



Machine Learning Pipeline



2. Train lead score model based on combined features

Gradient Boosting

Website Content	<div></div>	<div></div>	Company name
Company linkages	<div></div>	<div></div>	Industry codes
Products	<div></div>	<div></div>	Company age
Partnerships	<div></div>	<div></div>	Company size
Global Reach	<div></div>	<div></div>	Revenue
Publications	<div></div>	<div></div>	Profit/loss
Images	<div></div>	<div></div>	Equity
Jobposts	<div></div>	<div></div>	Subsidiaries

Machine Learning Pipeline

LEARN

PREDICT

11M
companies



...

Learned Model



Individual lead score
for each company



94%



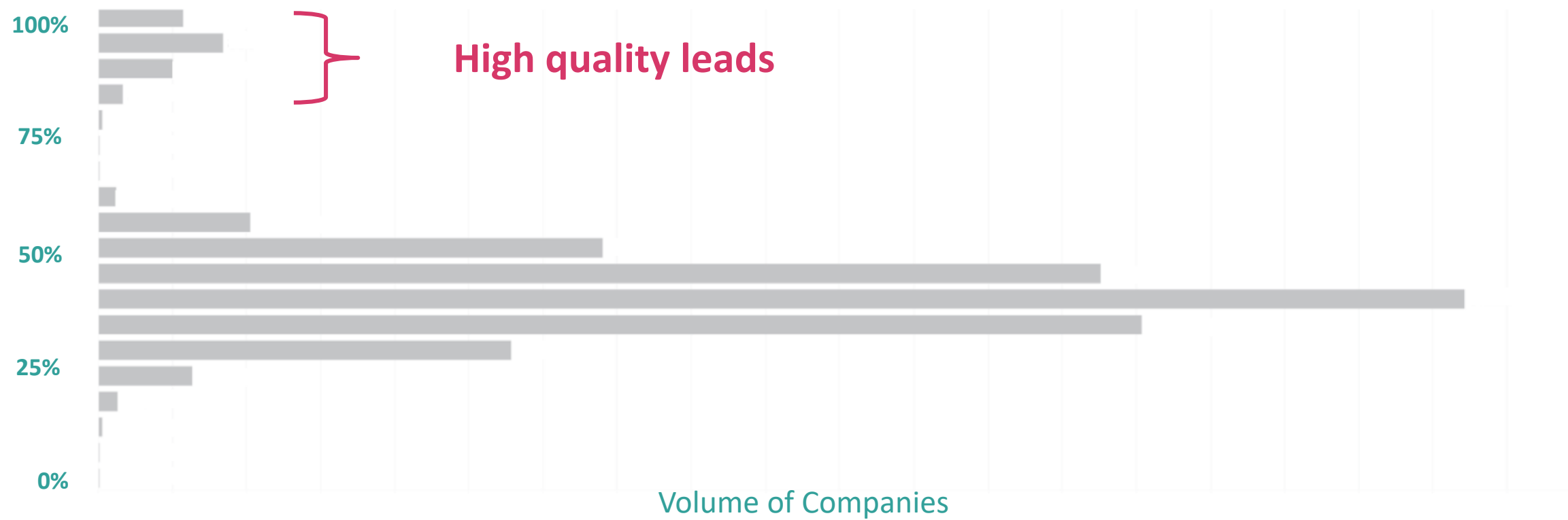
71%



33%

...

Results



Results



Heathcoat fabrics

Industry Code 2252

Establishments primarily engaged in manufacturing women's full-length and knee-length hosiery (except socks), and panty hose

Technical textile fair

Official industry clearly not relevant, but ...



Results



Textile
innovation

Search



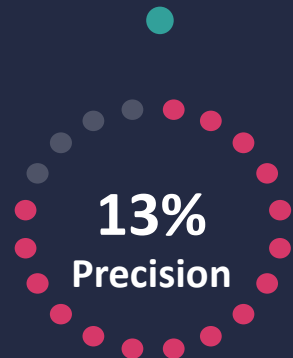
LATEST NEWS



Performances



Standard approach



Offline & Online features

