UNIVERSITÀ DI PISA

Machine Learning for Industry 4.0

Reshawn Ramjattan
Tutors: Prof. Daniele Mazzei & Prof. Davide Bacciu

Agenda

1. Introduction
2. Systematic Review
3. Routes for Future Work

Introduction

What is Industry 4.0?

The rapid changes due to increasing interconnectivity and smart automation

Types of ML applications?

- Predictive Maintenance
- Forecasting Load and Demand

Systematic Review

- There are reviews for top ML applications
- Not many of the entire intersection
- An infeasible number of papers to read
- Topic Modelling as an enabling technique

Systematic Review

Research Questions

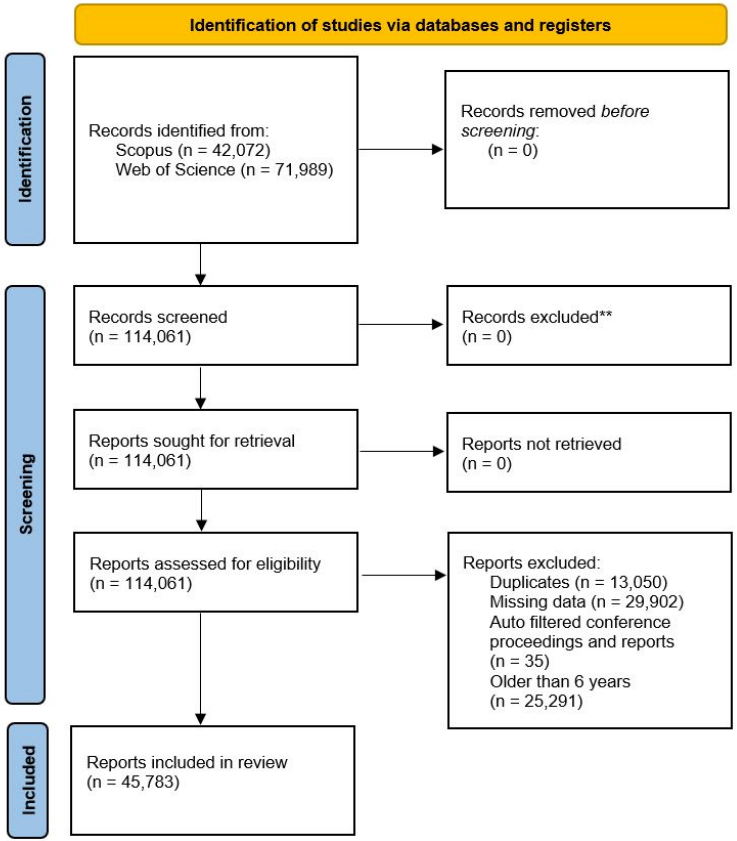
1. For what industry 4.0 problems are ML solutions most commonly discussed?
2. What ML methods see the most frequent use in these areas?
3. How do the areas focused on in academic literature compare to those in the whitepapers of top industrial companies?

Systematic Review

Paper Gathering

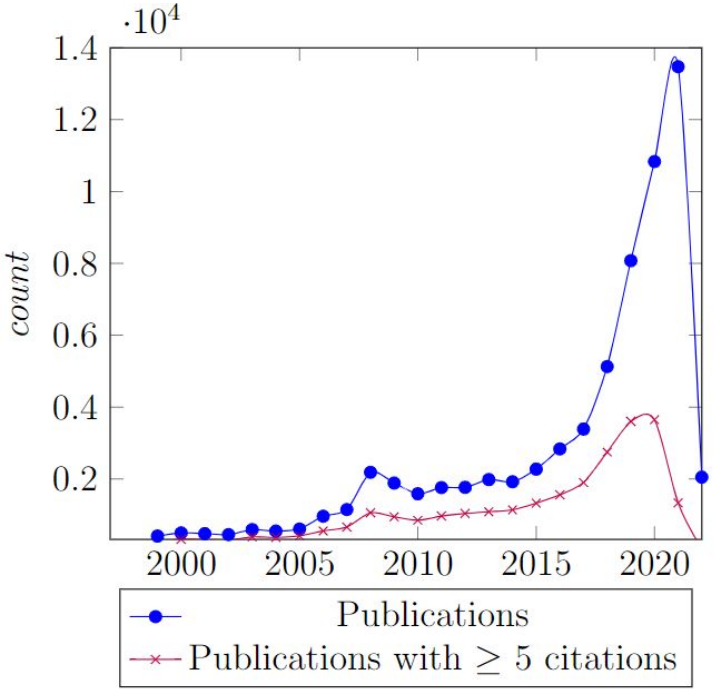
- Query: ("industrial" OR "industry 4.0")

  AND ("deep learning" OR "machine learning" OR "artificial intelligence" OR ….)
- Scopus, Web of Science
- Final set: 45,783 papers



Systematic Review

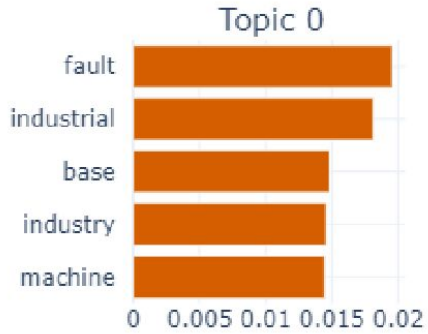
Papers over time



Systematic Review

Topic Modelling - BERTopic Algorithm

1. **Embeddings**: Get paper repr. with BERT using pre-trained model
2. **Reduce**: Dimensionality reduction with UMAP for better cluster results
3. **Cluster**: Cluster embeddings using HDBSCAN
4. **Topic Words**: Use variant of TF-IDF to get top words for each cluster

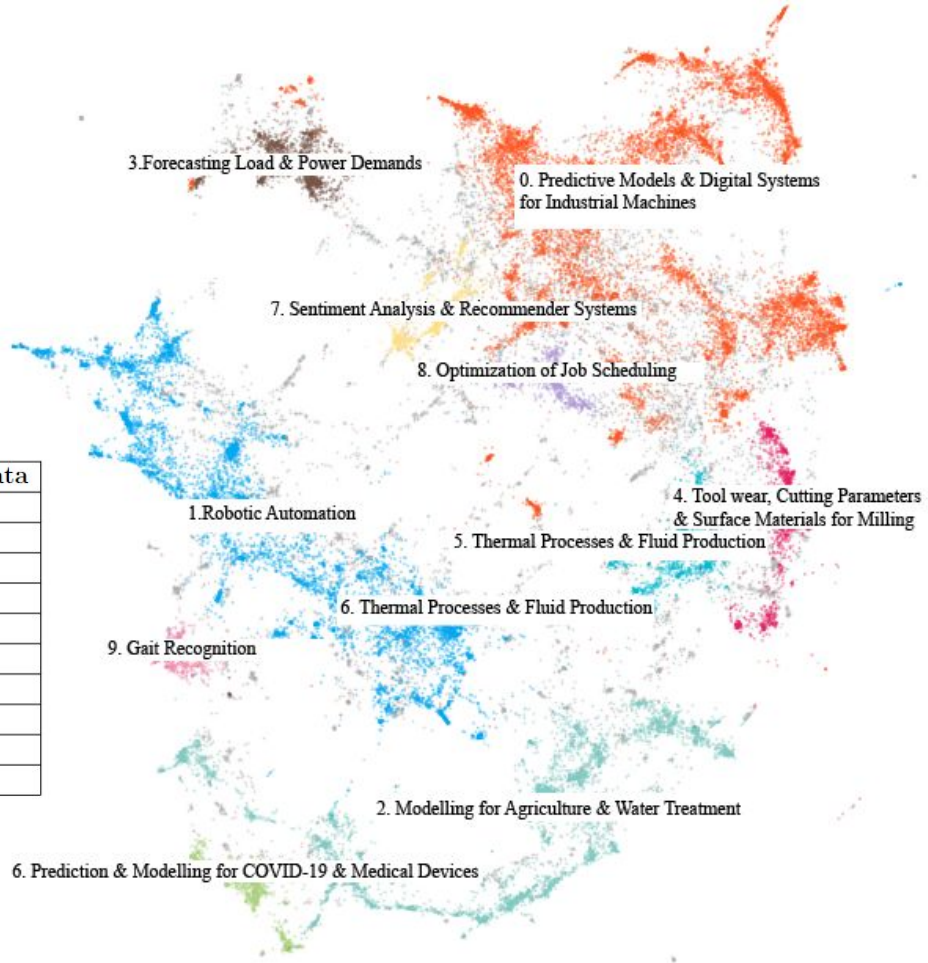


Systematic Review

Topic Modelling - Results

Table 2: Top 10 Results of Topic Modelling

Topic	% of Data
0. Predictive Models & Digital Systems for Industrial Machines	25.44
1. Robotic Automation	16.33
2. Modelling for Agriculture & Water Treatment	10.45
3. Forecasting Load & Power Demands	3.25
4. Tool wear, Cutting Parameters & Surface Materials for Milling	3.01
5. Prediction & Modelling of Thermic Processes & Fluid Production	2.56
6. Prediction & Modelling for COVID-19 & Medical Devices	2.37
7. Sentiment Analysis & Recommender Systems	2.37
8. Optimization of Job Scheduling	1.91
9. Gait Recognition & Modelling using Sensors, Wearables & Images	1.30



Systematic Review

What Industry problems are commonly discussed?

- Security/Risk Control
- Smart Production
- Connectivity
- Service Optimization
- Human-Machine-Interaction
- Supply Chain Management

Table 3: Top 20 Results of Topic Modelling inclusive of Sub-Topics

Topic	% of Data
0. Security & Intrusion Detection	4.38
1. Fault Diagnosis & Detection	3.88
2. Forecasting Load & Power Demands	3.25
3. Industrial IoT & Wireless Communication	3.05
4. Tool wear, Cutting Parameters & Surface Materials for Milling	3.01
5. Prediction & Modelling of Thermal Processes & Fluid Production	2.56
6. Prediction & Modelling for COVID-19 & Medical Devices	2.37
7. Sentiment Analysis & Recommender Systems	2.37
8. Reinforcement Learning for Robotics in Assembly & Planning	2.00
9. Optimization of Job Scheduling	1.91
10. Modelling Control of Robotic Manipulators	1.83
11. Modelling for Processing Wastewater & Water Treatment	1.80
12. Predictive Maintenance & RUL Forecasting	1.63
13. Adaptive Motor Control	1.60
14. Quality Inspection	1.37
15. Fuzzy Risk Analysis & Safety Management	1.36
16. Pose Estimation for Grasping Objects	1.33
17. Gait Recognition & Modelling using Sensors, Wearables & Images	1.30
18. Engineering Education & Skill Development	1.16
19. Forecasting Crop Yields	1.03

Systematic Review

ML Methods

- CNNs are the most common across the top 3
- RNNs see the most use in sequence problems
- RL has good presence despite being relatively less established in practical cases

Systematic Review

Perspective of Top Companies

- 17 recent and relevant Whitepapers from top companies (McKinsey, Accenture, Microsoft, Bain & Co., Deloitte, PwC, Boston Consulting Group)
- Manually reviewed for “high potential value” areas

Systematic Review

Mind Map of Topics

- Smart Production
- Connectivity
- Supply Chain Management
- Human-Machine-Interaction
- Data Analytics

Most common: Service Opt.

Most cost decreasing: Smart Prod.

Systematic Review

How do the topics compare to industry focus?

- Generally aligns well
- Focus on smart production cases but its justifiable
- Some topic areas just aren’t as “publishable”
- Companies are more focused on the bigger picture, not just the ML use case
- More future work on the topics that support adoption would be worthwhile

Going Forward

Democratized Technology

- Increasing access to technology
- Making it easier for non-experts to apply ML to industry problems
- Motivating adoption by demonstrating value as quickly as possible

Going Forward

Continual Learning

- Helps models adapt to changes that worsen performance
- While minimizing catastrophic forgetting
- Avoids complete retraining

Going Forward

Distributed ML

- Low-latency decentralized usage of models
- Ideal for low-bandwidth environments
- Security, privacy & reliability

Conclusion

- ML has a lot of potential in industry 4.0
- Security, Smart Prod., IoT, Service Optimization, Robotic Automation & Logistics
- CNNs were the most common ML method
- Academic lit. aligns with industry perspective
- Focusing more on integration and ease-of-use can help adoption

Thanks!

reshawn.ramjattan@phd.unipi.it