Application of Data mining in logistics in the Transition Era to Industry 4.0: Review

Authors: Khaoula Sahraoui, Samia Aitouche, Karima Aksa,
sahraouikhaoula12@gmail.com, s.aitouche@univ-batna2.dz, k.aksa@univ-batna2.dz
Plan

Introduction
Background
  industry 4,0
data mining
Methodology
  Descriptive analysis
Content analysis
Discussion and challenges
Conclusion
Introduction
In industry, precisely logistics, the tremendous volume of data in the industrial environment was increased, therefore, the use of data-mining as an analytical tool was necessary. The concept of big data and considerations of how to deal with such huge amount of data is an intrinsic challenge of any system operating in an industry 4.0 environment.

Our research aimed at revealing the location of research and researchers in using datamining for industry 4.0 in logistics and industry field in general.
Background

Industry

INDUSTRY 1.0
Mechanization, steam power, weaving loom
1784

INDUSTRY 2.0
Mass production, assembly line, electrical energy
1870

INDUSTRY 3.0
Automation, computers and electronics
1969

INDUSTRY 4.0
Cyber Physical Systems, internet of things, networks
TODAY
Bayes Theorem (1763)
Regression (1805)
Turing (1936)
Neural Networks (1943)
Evolutionary Computation (1965)
Databases (1970)
Genetic Algorithms (1975)
KDD (1989)
SVM (1992)
Data Science (2001)
MoneyBall (2003)
Big Data
Widespread adoption
DJ Patil (2015)
Methodology

In recognition of the recent and relevant research on datamining, industry 4.0 and logistics, a complex survey was conducted on January 19, 2020 in the databases: Scopus. The survey was also complex in order to advance relevance and limit the number of documents. It touches industry 4.0, datamining and logistics in titles and keywords.
Descriptive analysis

Analyze of the number of papers by year

- 2016: 3.57%
- 2017: 7.14%
- 2018: 28.57%
- 2019: 53.57%
- 2020: 7.14%
Analyze of the number of documents by country.
Descriptive analysis

Analyze of the number of papers by subject area

- Computer Science: 30.6%
- Engineering: 22.6%
- Mathematics: 11.3%
- Energy: 6.5%
- Business, Management, Accounting: 4.8%
- Decision Sciences: 4.8%
- Earth and Planetary Sciences: 4.8%
- Other: 3.2%
### Analyze of the keywords frequencies

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Frequency</th>
<th>Keywords</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry 4.0</td>
<td>19</td>
<td>Manufacturing</td>
<td>3</td>
</tr>
<tr>
<td>Enterprise</td>
<td>3</td>
<td>Supply chain</td>
<td>2</td>
</tr>
<tr>
<td>Manufacture</td>
<td>9</td>
<td>Industrial</td>
<td>4</td>
</tr>
<tr>
<td>Business</td>
<td>1</td>
<td>Data</td>
<td>14</td>
</tr>
<tr>
<td>Management</td>
<td>2</td>
<td>Machine learning</td>
<td>7</td>
</tr>
<tr>
<td>IoT or internet of things</td>
<td>4</td>
<td>Artificial intelligence</td>
<td>1</td>
</tr>
<tr>
<td>Data mining</td>
<td>5</td>
<td>Analytics + analysis</td>
<td>5</td>
</tr>
</tbody>
</table>

#### Keyword frequencies

- Active Management
- Waste-to-energy
- Sustainable Development
- Semantics
- Predictive Maintenance
- Manufacturing Plant
- Learning Algorithms
- Flue Gases
- Economic And Social Effects
- Data Integration
- Computer Vision
- State Of The Art
- Information Management
- Data Analytics
- Machine Learning
- Industry 4.0

Descriptive analysis
Content analysis

2016
(Sakornsathien K, 2016)
(Frischer R et al, 2017)
(Gokalp M. et al, 2017)

2017

2018
(DeMilia G et al, 2018),
(Villalobos K., 2018),
(Sternberg E. and Atzmudder M., 2018),
(Molano J.I.R.et al, 2018),
(Chaudhary K. et al, 2018),
(Rodríguez-Molano J.I. et al, 2018)

2019
(Jurdziak L. et al, 2019),
(Kebisek M et al, 2019),
(Penumuru D. et al, 2019),
(Anuşlu M.D. et al, 2019),
(Usuga Ca david, J. P. et al, 2020)
(Kabugo J. C et al, 2020)

2020
(Kebisek M et al, 2019),
(Penumuru D. et al, 2019),
(Anuşlu M.D. et al, 2019),
(Usuga Ca david, J. P. et al, 2020)
(Kabugo J. C et al, 2020)
Discussion and challenges

After analyzing the selected documents, a remarkable attentiveness head towards using datamining technics aiming at implement, develop and improve I4.0 in the industrial field generally or even know barriers and the constraints in front of this objective, and it is increasing from year to another.

Due to the complexity of industry 4.0 with the variant type and volume of data even the uncertainty, it is well known that simple use of DM algorithms will do not produce good results.

It is difficult to generalize the integrated data mining in all the technologies of the industry 4.0; it has to be very complicate and efficient algorithms and methods in order to catch the desired results using a high quality and strong potency machines that would be expansive.

Even though the industry 4.0 faces different challenges and difficulties involving many aspects, including scientific, technological, and economic challenges, in addition to social problems and political issues.
The group of documents that was chosen for this research, was the result of entering a query which limited the search to our domain target, however, many other articles have discussed this subject which could be analyze and treat. The content of the analyzed documents showed the rapidly spread of using datamining in industry 4.0, improving by that the supply chain management and logistics industry.

the improvement quality of the technologies that used in logistics and industry was noticed, concluding that data mining would give logistics and industry 4.0 a leap forward. Future research may focus on dealing with the faced challenges, also other challenges could appear by consulting other articles in this field.
Thank You for your attention
References


Frischer, Robert; Grycz, Ondrej; Hlavica, Robert: Concept Industry 4.0 in Metallurgical Engineering. 26th International Conference on Metallurgy and Materials (METAL), Brno, Czech Republic, May 24-26 (2017).


