



# **Production2Vec: a hybrid recommender system combining semantic and product complexity approach to improve industrial resiliency**

Arnault Pachot, Adélaïde Albouy-Kissi, Benjamin Albouy-Kissi, Frédéric Chausse

## **► To cite this version:**

Arnault Pachot, Adélaïde Albouy-Kissi, Benjamin Albouy-Kissi, Frédéric Chausse. Production2Vec: a hybrid recommender system combining semantic and product complexity approach to improve industrial resiliency. 2021 2nd International Conference on Artificial Intelligence and Information Systems (ICAIIIS '21), May 2021, Chongqing, China. 10.1145/3469213.3469218 . hal-03276942

**HAL Id: hal-03276942**

**<https://hal.science/hal-03276942>**

Submitted on 2 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Production2Vec: a hybrid recommender system combining semantic and product complexity approach to improve industrial resiliency

Arnault Pachot

Université Clermont-Auvergne, CNRS, SIGMA Clermont,  
Institut Pascal, F-63000 Clermont-Ferrand, France,  
arnault.pachot@etu.uca.fr, Corresponding author

Benjamin Albouy-Kissi

Université Clermont-Auvergne, CNRS, SIGMA Clermont,  
Institut Pascal, F-63000 Clermont-Ferrand, France  
benjamin.albouy@uca.fr

Adélaïde Albouy-Kissi

Université Clermont-Auvergne, CNRS, SIGMA Clermont,  
Institut Pascal, F-63000 Clermont-Ferrand, France  
adelaide.kissi@uca.fr

Frédéric Chausse

Université Clermont-Auvergne, CNRS, SIGMA Clermont,  
Institut Pascal, F-63000 Clermont-Ferrand, France  
frederic.chausse@uca.fr

## ABSTRACT

COVID-19 health crisis highlights the fragility of European industrial strategies and leads us to develop more agile, distributed, and resilient production models at a territorial level. There are two major challenges in this regard: one is to find solutions to secure supplies and/or industrial value chains, and the other is to identify companies that have the potential to transform their production quickly to cope with an emergency situation. We extended the Word2Vec vector space with products and economic activities allowing us to calculate proximities. We present a methodology based on semantic proximity and productive complexities to assess the ability of an A-company to produce a product B and to anticipate customer/supplier-type collaboration according to industrial quality standards. We consider recommendation topics by intertwining machine learning techniques with semantic approaches, referring to area ontologies incorporating territorial dimensionality.

## CCS CONCEPTS

• Natural language processing; • Recommender systems; • Supply chain management;

## KEYWORDS

Industrial resiliency, Recommender system

## ACM Reference Format:

Arnault Pachot, Adélaïde Albouy-Kissi, Benjamin Albouy-Kissi, and Frédéric Chausse. 2021. Production2Vec: a hybrid recommender system combining semantic and product complexity approach to improve industrial resiliency. In *2021 2nd International Conference on Artificial Intelligence and Information Systems (ICAIS '21)*, May 28–30, 2021, Chongqing, China. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3469213.3469218>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

ICAIS '21, May 28–30, 2021, Chongqing, China

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-9020-0/21/05...\$15.00

<https://doi.org/10.1145/3469213.3469218>

## 1 INTRODUCTION

The sudden disruption of supplies in March 2020, halted the production of many European companies, and the whole ecosystem that is linked to their activities. Facing this unprecedented situation, companies sought local solutions to source supplies, becoming aware of the importance of securing and diversifying their supplies.

In this article we adapt a word2vec based technique to a heterogeneous dataset of economic nomenclatures. We construct a productive vector space allowing us to perform productive recommendations between production units close to each other. From the product space, the recommender system is extended to identify production units whose production would potentially match if they make “productive jumps” to adapt to the needs.

The article is structured as follows. After a brief reminder of previous work on Word2Vec, the method and vector space are presented. This is then validated before using it to experiment a hybrid recommender system and results are presented on French area.

## 2 PREVIOUS WORK

Although currently NLP developments mainly implement Transformer [1] [2], the computation cost and the need for huge datasets leads to the use of different methods [3] [4] [5] [6] [7], which are based on the semantic proximity between words. Our work implements the well-established Word2Vec word embedding technique as presented in Figure 1. Word embedding consists of a distributed representation of each word in the form of a vector. Learning is done from specialized neural networks [8] or similar [9] [10] [11]. The Word2Vec model can be used in two variants to make a prediction of context (skip-gram) or word (CBOW). In addition to capturing syntactic and semantic information, the vectors produced by Word2Vec have geometric properties [11], which allow to represent larger blocks of information (such as sentences and paragraphs) by combining the dipping of words contained in the block. Word2Vec has also shown amazing ability to capture word concepts offering the capability to perform translators from simple linear transformations [12].

Word2Vec is used in many areas such as analysis of semantic proximity [13], trend analysis [14], predicting drug interactions [15] and sentiment analysis [6] [7]. Closer to our work, Word2Vec was

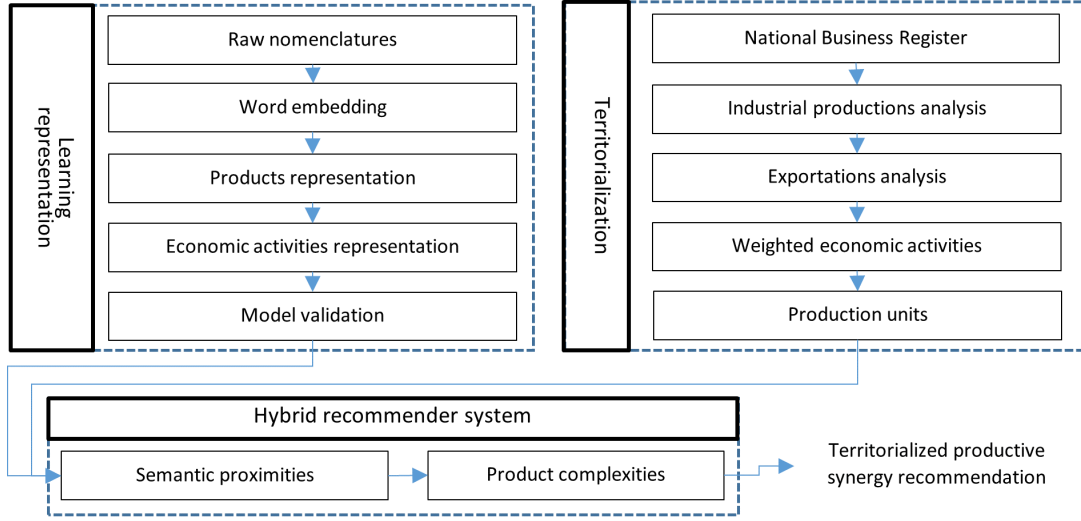


Figure 1: Production2Vec Architecture

Table 1: Nomenclatures

Nomenclature	Version	Type	# Classes	Publisher
HS 2017 <sup>a</sup>	2017	Product	5387	World Customs Organization
PRODCOM <sup>b</sup>	2019	Product	3919	European Commission
CPC <sup>a</sup>	2.1	Product	2887	United Nations
CPA <sup>b</sup>	2.1	Product	3218	European Commission
NC <sup>b</sup>	2020	Product	9484	European Commission
NACE <sup>b</sup>	2	Economic activity	615	European Commission
ISIC <sup>a</sup>	4	Economic activity	419	United Nations

<sup>a</sup> <https://unstats.un.org/unsd/classifications/Econ> <sup>b</sup> <https://ec.europa.eu/eurostat/ramon/index.cfm>

used to calculate similarities in industrial waste nomenclatures [16], taxonomy analysis [17] or to make recommendations [18] [19].

### 3 BUILDING A PRODUCTION VECTOR SPACE

#### 3.1 Word embedding from nomenclatures

Many nomenclatures are related to production. There are two main types of classifications: products and economic activities. Nomenclatures regularly evolve because of two main reasons: structural changes in the economy and standardization processes. Switching from one classification system to another is made by using correspondence tables.

As the nomenclatures are intended to standardize and structure information, special care has been taken in drafting class descriptions to limit ambiguities. This makes them particularly suitable for automated semantic processing. In our study we deal with the following nomenclatures, in their latest version.

We define a 100-dimensional vector space named  $\epsilon$  in which we will construct using Word2Vec a distributed representation of the words that are in the nomenclatures. We prepare a file containing 25929 records corresponding to the descriptions of every nomenclature

class described in table 1. We use the Skip-gram method[10] which at a nomenclature class (formed of a sequence of words)  $w_1, \dots, w_n$  seeks to maximize the average logarithmic probability of predicting the words of the sequence in relation to each other. Let  $k$  be the size of the training window and  $n$  the sequence size. The quantity to maximize is then defined as:

$$\frac{1}{n} \sum_{i=1}^n \left[ \sum_{j=-k}^k \log p(w_{i+j} | w_i) \right] \quad (1)$$

In this model every word  $w$  is associated with two learnable parameter vectors,  $u_w$  and  $v_w$ . The probability of predicting  $w_i$  knowing  $w_j$  is calculated from  $u_w$  and  $v_w$ , in a softmax function, with  $V$  representing the vocabulary size:

$$p(w_i | w_j) = \frac{e^{u_{w_i} \cdot v_{w_j}}}{\sum_{l=0}^V e^{u_l \cdot v_{w_j}}} \quad (2)$$

This yields a minimization problem, solved by hierarchical softmax[20]. The words will thus be represented by considering their proximity in the classes of nomenclatures, which enables us to capture semantic information.

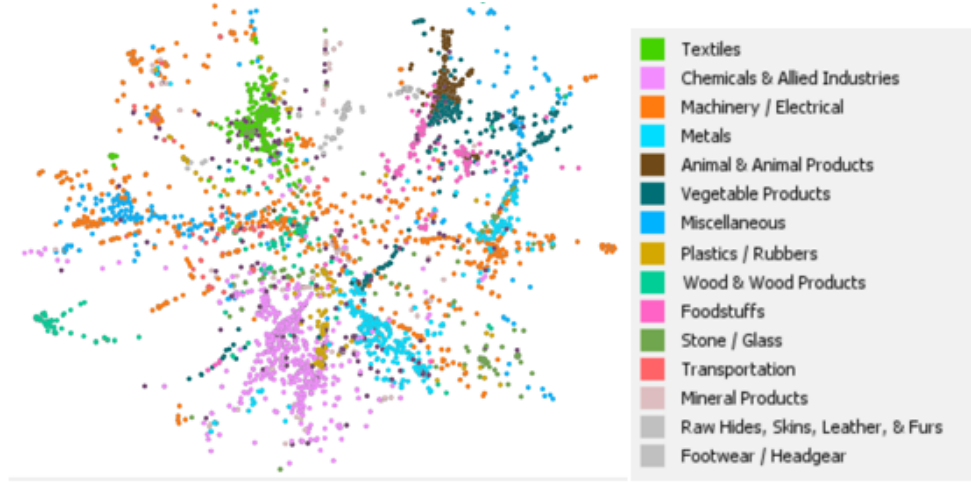


Figure 2: 2D Representation of products in the vector space. Products are colored according to their industry sector

### 3.2 Products representation

We choose the HS nomenclature as the reference nomenclature for the products. Consequently, for every 5387 classes of this nomenclature we will calculate a vector  $v_p$  representing the concept of the product  $p$  in the vector space  $E$ . Vector  $v_p$  is built from each word vector that appears in its class description in the HS nomenclature. The vector  $v_p$  is calculated by averaging the vectors of each word  $w_i$  that its class description contains.

### 3.3 Modeling economic activities

By resorting to a combination of correspondence tables between activities and products (NACE→CPA→HS)<sup>1</sup>, each NACE class is represented in the vector space  $E$  based on products it includes. We define economic activity  $a$  as an average vector of products from the HS nomenclature:  $v_a = \frac{1}{n} \sum_{i=1}^n v_{p_i}$  with  $p_1, \dots, p_n$  the products associated with  $a$ .

**3.3.1 Extending the model using a weighted correspondence table.** Since 1995, the European Commission has published annual production statistics by country and by product type called PRODCOM<sup>2</sup>. For 2617 types of industrial products, the volumes and the amount of production are provided. We then convert them into the HS nomenclature. Using the correspondence table between activities and products we calculate a weight for every product class that is related to an activity, basing ourselves on production amounts per product. Let  $p_1, \dots, p_n$  be the list of products associated with economic activity  $a$  and the amount of production  $prod(p_i)$  of the product  $p_i$ , then the weight  $\gamma_{p_i}$  of the product  $p_i$  within the activity

$a$  is defined as follows:

$$\gamma_{p_i} = \frac{prod(p_i)}{\sum_{j=0}^n prod(p_j)} \quad (3)$$

Using the same method, we processed export data (export amount per product). The advantage of this alternative lies in the fact that the data cover more classes of products (even non-industrial products). Moreover, territorialized data are available for some countries, such as France where Customs Service provides export volumes information for every department<sup>3</sup>. We can then weight productions knowing how territories consider their productive distinctiveness. Following the analysis presented above, we also define the vector associated with economic activity in a productive area  $\tau$ , as the weighted sum of the vectors of the products associated with it in  $\tau$ . With  $p_1, \dots, p_n$  the products and  $\lambda_1, \dots, \lambda_n$  their respective weights associated with an activity  $a$  in a productive area  $\tau$ , compute the vector  $v_a(\tau) = \sum_{i=1}^n \lambda_i v_{p_i}$ , with  $\sum_{i=0}^n \lambda_i = 1$

### 3.4 Representing production units

We relied on the SIRENE<sup>4</sup> dataset that is available history of French production units since 1973. It provides information for every facility, relating them to their connected company, their economic sector, their workforce group, and their postal address. By using a geolocation service, we get a GPS position for each facility allowing us to select production units located on a productive area. From the headcount groups we calculate an average workforce. We also make up an overall index of the cumulative average workforce by industry, year, and department.

We represent production units basing ourselves on their productive capacity, which is defined as production in relation with their workforce-weighted industry. Let  $\alpha$  be the workforce-weight, the

<sup>1</sup> We parse the NAF to CPF correspondence document ([https://www.insee.fr/fr/statistiques/fichier/2399243/Nomenclatures\\_NAF\\_et\\_CPF\\_Reedition\\_2020.pdf](https://www.insee.fr/fr/statistiques/fichier/2399243/Nomenclatures_NAF_et_CPF_Reedition_2020.pdf)) to obtain a NACE 2 to CPF 2.1 CSV file. NACE is obtained from NAF code by removing the last letter. CPF and CPA are identical in version 2.1. Matching between CPA 2.1 and HS 2017 is done using the CPA 2.1 to NC 2017 correspondence table ([https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST\\_REL](https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL)) because HS code equivalent to the sixth first digits of the related NC code.

<sup>2</sup> <https://ec.europa.eu/eurostat/web/prodcom/data/excel-files-nace-rev.2>

<sup>3</sup> <https://www.data.gouv.fr/fr/datasets/statistiques-regionales-et-departementales-du-commerce-exterieur/>

<sup>4</sup> <https://www.sirene.fr/sirene/public/static/acces-donnees>

**Table 2: Top@n accuracy for NACE activity (615 classes)**

Corpus	Vector size	Vocab. size	Top-1	Top-3	Top-5	Top-10
<b>Nomenclatures</b>	<b>100</b>	<b>12234</b>	<b>78.6%</b>	<b>89.0%</b>	<b>92.7%</b>	<b>94.1%</b>
English Wikipedia Dump of February 2017	300	302866	72.7%	86.8%	90.5%	93.7%
Gigaword 5th Edition	300	292479	72.0%	84.9%	88.9%	91.5%
English CoNLL17 corpus	100	4027169	73.1%	84.9%	88.6%	91.9%

products  $p_1, \dots, p_n$ , and their respective  $\lambda_1, \dots, \lambda_n$  weights associated with production units  $e$ . With  $\sum_{i=0}^n \lambda_i = 1$ , vector  $v_e$  is:

$$v_e = \alpha \sum_{i=1}^n \lambda_i v_{p_i} \quad (4)$$

## 4 EXPERIMENTATION

### 4.1 Productive space validation

We predict activity class using two different methods (semantic method by word embedding and official correspondence table) and compare the results:

- The economic activities' vectors that were designed by calculating the average vector of every word contained in their class descriptions.
- The economic activities' vectors of our model that were designed by using a correspondence table (see section 4.3).

Because nomenclatures are taxonomies, we can perform the comparison by conducting a hierarchical search, which means that, at first, we looked for the closest category (comparing in pairs every category vector, built by average vector of each class they contain) and then looking inside this category, the product vector that is the most similar. The results are presented in the table below. We have also carried out several tests with different corpuses that show that the approach we have chosen (learning on a corpus made up of nomenclatures) provides the best results compared to the results obtained on a pre-trained model. Classes are predicted with a top-1 accuracy of 78.6%.

### 4.2 Exploiting the vector space to recommend productive collaborations between production units

**4.2.1 Measuring productive synergies using semantic proximities.** The distance between economic activities is calculated from the distance between the products associated with each economic activity. Productive distance between economic activities  $a$  and  $b$  over the productive area  $\tau$  is defined as:

$$\Psi_{a,b}(\tau) = \frac{1}{n \cdot m} \sum_{i=0}^n \sum_{j=0}^m \left( \lambda_{a_i} \lambda_{b_j} \frac{v_{p_{a_i}} \cdot v_{p_{b_j}}}{\|v_{p_{a_i}}\| \|v_{p_{b_j}}\|} \right) \quad (5)$$

With  $v_p$  the vector associated with the product  $p$  in  $E$ ,  $a$  and  $b$  two economic activities in a productive area  $\tau$ ,  $p_{a_1}, \dots, p_{a_n}$  and  $\lambda_{a_1}, \dots, \lambda_{a_n}$  respectively the products and their weights associated with the activity  $a$ ,  $p_{b_1}, \dots, p_{b_m}$  and  $\lambda_{b_1}, \dots, \lambda_{b_m}$  respectively the products and their weights associated with the activity to  $b$ .

Looking for potential partners of production unit  $e$  in a productive area  $\tau$  is similar to looking production units whose economic activity is close to  $e$ . We calculate the recommendations proposed on the production unit  $e$  of a given productive area  $\tau$  with an economic activity  $a$  by building a list of productive units  $\{g_1, \dots, g_n\}$ , with economic activities respectively  $\{b_1, \dots, b_n\}$  ordered inversely to  $\Psi_{a,b_i}(\tau)$  and a threshold  $\theta$ .

$$\hat{R}_{e,\tau} = \{g_1, \dots, g_n | \Psi_{a,b_i}(\tau) \geq \Psi_{a,b_{i+1}}(\tau) \text{ and } \Psi_{a,b_i}(\tau) > \theta\} \quad (6)$$

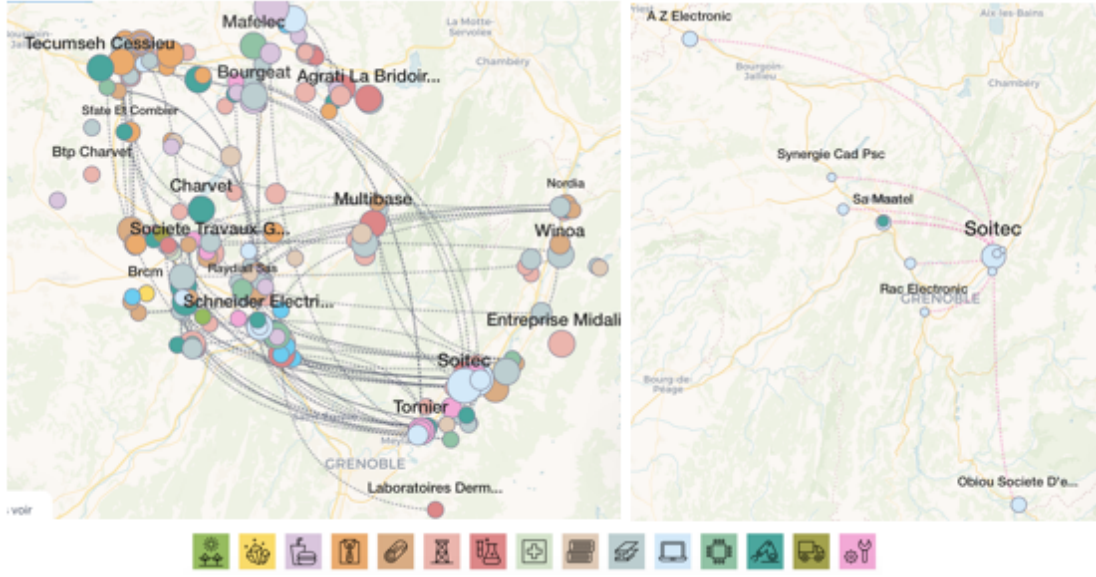
**4.2.2 Identifying productive jumps and hybridation with productive synergies.** A territory's ability to adapt its production tool in the event of a systemic shock to produce the necessary products contributes to its resilience[21][22]. We have seen such adaptations during the COVID19 crisis when spirits or paint companies were able to produce hydro-alcoholic gel as a matter of urgency. We can estimate the potential "productive jumps" of a production unit, based on the work done by Harvard University[23][24][25] on the economic complexity of productive systems. This work provides a product space[26] where we can measure proximity of know-how between two product classes. A closeness of know-how between the product  $p_1$  and the product  $p_2$  indicates that a production facility capable of producing  $p_1$  could produce  $p_2$  with limited adaptation of its production tool. Let  $X_{c,p}$  be the value of exports in country  $c$  for the product  $p$ . Revealed Competitive Advantage  $RCA_{cp}$  greater than or equal to 1 indicates that  $c$  export product  $p$ .  $RCA_{cp}$  less than 1 indicates that country  $c$  is not an effective exporter of  $p$ .

$$RCA_{cp} = \left( \frac{X_{cp}}{\sum_c X_{cp}} \right) / \left( \frac{\sum_p X_{cp}}{\sum_{c,p} X_{cp}} \right) \quad (7)$$

With  $M_{cp} = \begin{cases} 1 & \text{if } RCA_{cp} \geq 1 \\ 0 & \text{else} \end{cases}$  indicating whether the country  $c$  has a Revealed Competitive Advantage on the product  $p$ , the level of productive proximity  $\phi_{p_1,p_2}$  between each product pair  $\{p_1, p_2\}$  in the product space is based on export data from countries using the following formula[25]:

$$\phi_{p_1,p_2} = \min \left( \frac{\sum_c M_{cp_1} M_{cp_2}}{\sum_c M_{cp_1}}, \frac{\sum_c M_{cp_1} M_{cp_2}}{\sum_c M_{cp_2}} \right) \quad (8)$$

Knowing the productive capabilities of a territory's production units, we can identify the territory's ability to manufacture new products. Let  $p'$  be the product obtained by "jumping" from  $p$  in the product space, such that it does not exist  $q$  giving  $\phi_{p,q} > \phi_{p,p'}$ . We reformulate the productive distance between two economic activities  $a$  and  $b$  over the productive area  $\tau$  by integrating proximity



**Figure 3: On left: visualization of potential productive synergies between companies (example is taken in France around the city of Grenoble). Nods colors corresponds to industry sectors. On right: example of personalized partners recommendation for the company “Soitec”, dedicated in the manufacture of electronic components (code NACE 21.11)**

in the product space:

$$\Psi'_{a,b}(\tau) = \frac{1}{n \cdot m} \sum_{i=0}^n \sum_{j=0}^m \left( \lambda_{a_i} \lambda_{b_j} \max \left( \frac{v_{p_{a_i}} \cdot v_{p_{b_j}}}{|v_{p_{a_i}}| |v_{p_{b_j}}|}, \frac{v_{p_{a_i}} \cdot v_{p'_{b_j}}}{|v_{p_{a_i}}| |v_{p'_{b_j}}|} \right) \right) \quad (9)$$

We calculate a recommendation  $\hat{R}'_{e,\tau}$  that includes productive jump in the same manner that the recommendation (6) but using  $\Psi'$  instead of  $\Psi$ .

Our recommendation system seeks to identify, for a given production unit  $a$ , production units located in a productive area whose economic activity is close to  $a$  (above a threshold). In case of failure, it is possible to propose an alternative: production units whose economic activity would be potentially close to  $a$  if they make productive jumps. We combine methods, using cascading mode: the second method of recommendation (productive jumps) will be used if the first method (semantic proximity) fails.

$$\widehat{RS}_{e,\tau} = \begin{cases} \hat{R}'_{e,\tau} & \text{if } |\hat{R}'_{e,\tau}| \neq 0 \\ \hat{R}_{e,\tau} & \text{else} \end{cases} \quad (10)$$

## 5 RESULTS

From French datasets, we built an interactive map of France’s productive know-how and potential industrial synergies. The recommender system offers the opportunity to seek out local productive partners to secure their supplies, and thus promote distributed manufacturing. The territorial data-based model showed interesting predictive capabilities to recommend industrial collaborations. Modelling the production of a territory combined with the knowledge of potential productive jumps also allows us to visualize the potential production of a territory: the goods that are currently

manufactured and those that could be manufactured in the event of a destabilizing shock.

## 6 CONCLUSION AND PERSPECTIVES

Semantic analysis of nomenclatures combined with products complexities opens up new perspectives to model territorial productive systems. Various applications are thus made possible: for instance, company’s supplies may then be secured by searching for local suppliers, by choosing a new settlement for a relocation, or by defining the prospects of companies’ diversification or co-production.

This methodology offers new perspectives to improve industrial production in delimited areas. We need to define an industrial resilience indicator that can be measured from modeled systems, based on the resilience of complex systems [21]. We will integrate the HR dimension (jobs and demographics) and waste to foster synergies between the players of the circular economy [16].

## REFERENCES

- [1] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171-4186. <https://doi.org/10.18653/v1/N19-1423>
- [2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>
- [3] Almeida, F., & Xexéo, G. (2019). Word Embeddings: A Survey. *arXiv:1901.09069 [cs, stat]*. <http://arxiv.org/abs/1901.09069>
- [4] Di Gennaro, G., Buonanno, A., Di Girolamo, A., Ospedale, A., & Palmieri, F. A. N. (2021). Intent Classification in Question-Answering Using LSTM Architectures. In A. Esposito, M. Faundez-Zanuy, F. C. Morabito, & E. Pasero (Eds.), *Progresses in Artificial Intelligence and Neural Systems* (p. 115-124). Springer Singapore. [https://doi.org/10.1007/978-981-15-5093-5\\_11](https://doi.org/10.1007/978-981-15-5093-5_11)



- [5] Jang, B., Kim, I., & Kim, J. W. (2019). Word2vec convolutional neural networks for classification of news articles and tweets. *PLOS ONE*, 14(8), e0220976. <https://doi.org/10.1371/journal.pone.0220976>
- [6] Jianqiang, Z., Xiaolin, G., & Xuejun, Z. (2018). Deep Convolution Neural Networks for Twitter Sentiment Analysis. *IEEE Access*, 6, 23253–23260. <https://doi.org/10.1109/ACCESS.2017.2776930>
- [7] Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews. *Procedia Computer Science*, 179, 728–735. <https://doi.org/10.1016/j.procs.2021.01.061>
- [8] Bengio, Y., Ducharme, R., Vincent, P., & Janvin, C. (2003). A Neural Probabilistic Language Model. *J. Mach. Learn. Res.*, 3(null), 1137–1155.
- [9] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural Language Processing (almost) from Scratch. *ArXiv:1103.0398 [Cs]*. <http://arxiv.org/abs/1103.0398>
- [10] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *ArXiv:1301.3781 [Cs]*. <http://arxiv.org/abs/1301.3781>
- [11] Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *ArXiv:1310.4546 [Cs, Stat]*. <http://arxiv.org/abs/1310.4546>
- [12] Mikolov, T., Le, Q. V., & Sutskever, I. (2013). Exploiting Similarities among Languages for Machine Translation. *ArXiv:1309.4168 [Cs]*. <http://arxiv.org/abs/1309.4168>
- [13] Sultan, M. A., Bethard, S., & Sumner, T. (2015). DLS@SCU: Sentence Similarity from Word Alignment and Semantic Vector Composition. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 148–153. <https://doi.org/10.18653/v1/S15-2027>
- [14] sKim, S., Park, H., & Lee, J. (2020). Word2vec-based latent semantic analysis (W2V-LSA) for topic modeling: A study on blockchain technology trend analysis. *Expert Systems with Applications*, 152, 113401. <https://doi.org/10.1016/j.eswa.2020.113401>
- [15] Zhang, Y.-F., Wang, X., Kaushik, A. C., Chu, Y., Shan, X., Zhao, M.-Z., Xu, Q., & Wei, D.-Q. (2020). SPVec: A Word2vec-Inspired Feature Representation Method for Drug-Target Interaction Prediction. *Frontiers in Chemistry*, 7, 895. <https://doi.org/10.3389/fchem.2019.00895>
- [16] van Capelleveen, G., Amrit, C., Zijm, H., Yazan, D. M., & Abdi, A. (2021). Toward building recommender systems for the circular economy: Exploring the perils of the European Waste Catalogue. *Journal of Environmental Management*, 277, 111430. <https://doi.org/10.1016/j.jenvman.2020.111430>
- [17] Eljasik-Swoboda, T., Hemmje, M., Dascalu, M., & Trausan-Matu, S. (2016). *Combining Taxonomies using Word2vec*. 131–134. <https://doi.org/10.1145/2960811.2967151>
- [18] Baek, J.-W., & Chung, K.-Y. (2020). Multimedia recommendation using Word2Vec-based social relationship mining. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-019-08607-9>
- [19] Esmeli, R., Bader-El-Den, M., & Abdullahi, H. (2020). Using Word2Vec Recommendation for Improved Purchase Prediction. *2020 International Joint Conference on Neural Networks (IJCNN)*, 1–8. <https://doi.org/10.1109/IJCNN48605.2020.9206871>
- [20] Morin, F., & Bengio, Y. (2005). Hierarchical Probabilistic Neural Network Language Model. *AISTATS*.
- [21] Fraccascia, L., Giannoccaro, I., & Albino, V. (2018). Resilience of Complex Systems: State of the Art and Directions for Future Research. *Complexity*, 2018, 1–44. <https://doi.org/10.1155/2018/3421529>
- [22] Sheffi, Y., & Rice, J., Jr. (2005). A Supply Chain View of the Resilient Enterprise. *MIT Sloan Management Review*, 47.
- [23] Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26), 10570–10575. <https://doi.org/10.1073/pnas.0900943106>
- [24] Hausmann, R., & Hidalgo, C. (2010). Country Diversification, Product Ubiquity, and Economic Divergence. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1724722>
- [25] Hausmann, R., Hidalgo, C. A., & Bustos, S. (2013). *The atlas of economic complexity: Mapping paths to prosperity*. The MIT Press.
- [26] University, T. G. L. A. H. (2019). *Product Space Networks* [Data set]. Harvard Dataverse. <https://doi.org/10.7910/DVN/FCDZBN>