Enhancing Industrial Productivity through AI-Driven Systematic Literature Reviews

Jaqueline Gutierri Coelho[2,5,*][a], Guilherme Dantas Bispo[2,5][b], Guilherme Fay Vergara[1,2,8][c], Gabriela Mayumi Saiki[2,5][d], André Luiz Marques Serrano[1,2,4][e], Li Weigang[1,2,5,*][f], Clovis Neumann[1,2,4][g], Patricia Helena Martins[2,6][h], Welber Santos de Oliveira[1,2,8][i], Angela Brigida Albarello[2,5][j], Ricardo Accorsi Casonatto[2,4][k], Patrícia Missel[1,2][l], Roberto de Medeiros Junior[3,7][m], Jefferson de Oliveira Gomes[3,7][n], Carlos Rosano-Peña[9][o] and Caroline Cabral F. da Costa[3,7][p]

[1] *Latitude, Decision Making Technologies Lab, University of Brasilia, Campus Darcy Ribeiro, Brasilia, Brazil*

[2] *Projectum, Research Group for Math. Methodologies Applied to Management, University of Brasilia, Brasilia, Brazil*

[3] *SENAI, National Service of Industrial Learning, Federal District, Brazil*

[4] *Department of Production Engineering, Faculty of Technology, University of Brasília, Federal District, Brazil*

[5] *Department of Computer Science, University of Brasilia, Federal District, Brazil*

[6] *Department of Economics, University of Brasília, Federal District, Brazil*

[7] *UNITEC, Innovation, and Technology Unit, National Service of Industrial Learning, Federal District, Brazil*

[8] *Department of Electrical, University of Brasilia, Federal District, Brazil*

[9] *Department of Administration, University of Brasilia, Federal District, Brazil*

Keywords: Artificial Intelligence, Automatic Classifier, Innovation, Productivity in Industry, Sustainability, SLR.

Abstract: The advent of Artificial Intelligence (AI) has opened up new possibilities for improving productivity in various industry sectors. In this paper, we propose a novel framework aimed at optimizing systematic literature reviews (SLRs) for industrial productivity. By combining traditional keyword selection methods with AI-driven classification techniques, we streamline the review process, making it more efficient. Leveraging advanced natural language processing (NLP) approaches, we identify six key sectors for optimization, thereby reducing workload in less relevant areas and enhancing the efficiency of SLRs. This approach helps conserve valuable time and resources in scientific research. Additionally, we implemented four machine learning models for category classification, achieving an impressive accuracy rate of over 75%. The results of our analyses demonstrate a promising pathway for future automation and refinements to boost productivity in the industry.

1 INTRODUCTION

In a rapidly changing technological landscape, the efficiency of resource allocation and the preservation of knowledge capital are critical considerations. Within this context, productivity serves as a vital indicator, assessed primarily through labor productivity (LP) and total factor productivity (TFP). LP measures wealth generated per worker by dividing GDP by annual labor hours, while TFP gauges the comprehensive impact of inputs like labor development and capital assets in the production process. These measures are fundamental to assessing economic progress and

[a] https://orcid.org/0000-0002-6517-1957
[b] https://orcid.org/0000-0002-4938-2076
[c] https://orcid.org/0000-0002-4551-2240
[d] https://orcid.org/0009-0008-5941-1601
[e] https://orcid.org/0000-0001-5182-0496
[f] https://orcid.org/0000-0003-1826-1850
[g] https://orcid.org/0000-0003-4320-8795
[h] https://orcid.org/0000-0002-1511-6239
[i] https://orcid.org/0009-0002-7020-5442
[j] https://orcid.org/0009-0002-9455-0144
[k] https://orcid.org/0009-0001-3652-5525
[l] https://orcid.org/0009-0009-2435-4229
[m] https://orcid.org/0009-0002-7436-0554
[n] https://orcid.org/0000-0002-6004-799X
[o] https://orcid.org/0000-0002-6868-9284
[p] https://orcid.org/0000-0003-2005-3823

Coelho, J., Bispo, G., Vergara, G., Saiki, G., Serrano, A., Weigang, L., Neumann, C., Martins, P., Santos de Oliveira, W., Albarello, A., Casonatto, R., Missel, P., Medeiros Junior, R., Gomes, J., Rosano-Peña, C. and F. da Costa, C.
Enhancing Industrial Productivity Through AI-Driven Systematic Literature Reviews.
DOI: 10.5220/0012235000003584
In *Proceedings of the 19th International Conference on Web Information Systems and Technologies (WEBIST 2023)*, pages 472-479
ISBN: 978-989-758-672-9; ISSN: 2184-3252

competitiveness related to the evolution of the technology (Mundial, 2018).

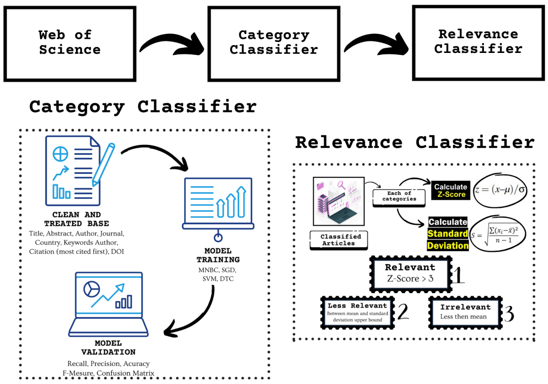

Figure 1: Framework of AI-Driven Systematic Literature Reviews.

To explore this direction further, the current study presents an AI-Driven Systematic Literature Reviews (SLR) framework, depicted in Figure 1, along with an optimized four-step method to yield the desired outcomes. The research relies on a publicly accessible web research database named Web of Science (WoS) to source the most relevant scientific articles across six predefined categories.

In the initial step, datasets are obtained from WoS using six distinct queries for each category. Subsequently, the combination of these datasets creates a labeled database that facilitates the application of various artificial intelligence algorithms to achieve the desired outcomes. The research not only utilizes these databases but also emphasizes the impact of each method employed. It conducts a critical and comparative analysis of these methodologies and discusses their practical implications. The primary objective is to gain insights into how different methods model the SLR process, the underlying phases involved, the assumptions guiding their application, and other contributing factors to this process. By exploring these aspects, the study seeks to contribute valuable knowledge and understanding enhancing the efficacy of systematic literature reviews in the context of labor productivity research.

2 LITERATURE REVIEW

This section aims to present how the literature provides the necessary information for the review of existing literature and practices on productivity. In recent years, studies on the evolution of productivity have increasingly gained ground in the economic debate in Brazil. This can be seen by the number of studies such as those by Torezani (2018), Santos et al. (2019), Borracini (2021), among others. Thus, understanding the pattern of productivity evolution is justified by the need to assimilate a country's competitiveness, either to maintain a space in the international scenario or to allow economic growth.

2.1 Systematic Literature Review (SLR)

Systematic literature review (SLR), a long-standing practice in health care, was consolidated in the late 1980s with the publication of the book Effective care during pregnancy and childbirth. In this context, it has been used as a way to overcome weaknesses and minimize biases left by narrative reviews, whose guiding threads can follow the multiple domains of knowledge arising from it. The following steps need to be followed: Subject search sources, Strategies for research bias, Evaluation of the studies and literature selected for SLR, Synthesis of results, and Study report.

It is noteworthy that one of the main strengths of the SLR is the focus on a specific search, the clarity in retrieving articles for review, objective and quantitative summaries, and evidence-based inferences Adesope et al. (2010). Galvão et al. (2017) summarize four types of mixed reviews: the quantitative convergence mixed review, the qualitative convergence mixed review, the exploratory sequential hybrid review, and the explanatory sequential mixed review. The quantitative convergence mixed review translates qualitative, quantitative, and mixed studies into qualitative findings. The exploratory sequential hybrid review occurs in two phases: transformation of results into qualitative findings and tabulation and comparison of quantitative results. The exploratory sequential mixed review measures the effects of an action and explains their differences.

Regarding the quality of reviews, the Oxford Classification of Level of Evidence established by Howick (2011), understands that systematic reviews of cross-sectional studies, systematic reviews of cohort studies, and systematic reviews of randomized controlled trials have level 1 evidence on a scale of 1 to 5, with 1 being the highest level of evidence. Therefore, this classification highlights the importance of systematic reviews for scientific progress and different decision-making.

On the other hand, Cohen et al. (2006), point out that the automatic classification of documents is an effective practice that brings clarity to the systematic literature review. Following this line, de Melo et al. (2022) demonstrate a change in the 2006 metric, now called Adjusted Work Saved over Sampling (AWSS@R), such metric brings the possibility of comparison between different domains, especially with regard to the automated citation screening step.

3 CATEGORY OF PRODUCTIVITY IN INDUSTRY

In our study, we conducted a systematic literature review by the traditional method of manual keyword selection and filtering. We then used the Category Classification of Scientific Articles and the Relevance Classification of Scientific Articles as a fair workload measure to compare and demonstrate its efficiency.

To conduct a comprehensive bibliometric review on the topic of Productivity in Industry, we will adopt a rigorous approach by applying specific filters to ensure the relevance and timeliness of the included studies. In order to cover a significant time range, we will analyze articles published in the period 2014 to 2023, considering several reliable sources.

Within the context of productivity in industry, there are six main axes that play essential roles in the advancement and transformation of the sector. These axes are innovation, sustainability, industry 4.0, industry 5.0, efficiency, and artificial intelligence. We will explore each of them below:

Innovation improves industrial productivity by introducing new ideas, technologies, and processes, optimizing operations, and boosting efficiency. It also allows for adaptation to market changes and the search for creative solutions to industrial challenges.

Sustainability is key in the industry, with practices that reduce resource consumption, minimize environmental impact and promote energy efficiency. Companies adopt sustainable strategies to improve productivity and reduce their environmental impact.

Industry 4.0 integrates advanced digital technologies such as IoT, big data, and artificial intelligence to create smart factories. This results in automation, improved quality control, and higher productivity.

Industry 5.0 integrates advanced technologies, such as collaborative robotics and artificial intelligence, with human labor to improve productivity. Humans and machines work together, leveraging their unique abilities, combining efficiency and creativity.

Efficiency Is fundamental to industrial productivity, involving process optimization, waste reduction, and quality improvement. It is achieved through continuous analysis and improvement of processes, aiming for more effective operational performance.

Artificial Intelligence (Weigang et al., 2022) drives industrial productivity through automation, supply chain optimization, and data-driven decision-making, resulting in greater efficiency and quality.

These six axes interact with each other, forming a set of approaches and technologies that contribute to boosting productivity in the industry. The adoption and integration of these elements into industrial strategies and processes can result in significant gains in efficiency, competitiveness, and sustainability.

4 SLR BY WEB OF SCIENCE

Web of Science by Clarivate Analytics is a widely recognized and extensively utilized platform among researchers, providing valuable access to academic resources, including articles and conferences, as well as a comprehensive citation index. Its robust features, such as research profiles and citation alerts, facilitate the exploration of relevant scholarly works and enable the monitoring of publication impact. Undoubtedly, this tool plays a crucial role in supporting systematic reviews, owing to its expansive coverage and advanced filtering capabilities. Consequently, Web of Science significantly contributes to the advancement of academic knowledge.

4.1 Database

Therefore, the Web of Science is essential for structuring the data of the articles for the analysis of each axis. To structure the article data for analysis in each category, the process involves:

1. Selection of categories: researchers begin by selecting the relevant categories pertinent to their analysis;
2. Keyword search in WoS: utilizing specific keywords, such as ”industry”, ”productivity”, and the chosen category names, it was possible to perform complete searches in Web of Science;
3. Creation of new column for category label: upon retrieving relevant data from the searches, researchers create new columns in their dataset to label each selected category;
4. Replication of the process for other categories: repeat the search and data retrieval process for the remaining five categories, creating dedicated columns for each;
5. Compilation of databases: the data obtained from the six categories are then compiled into a unified database, referred to as the ”Labeled Database”;
6. Filtering of Replications: to avoid redundancies,

duplicates were removed ensuring that each article is linked to only one category.

After completing the data extraction stage, a unified database was obtained, structured in nine columns: Title, Abstract, Author, Journal, Country, Keywords, Citation (ordered by frequency, with the most frequent citations at the top), DOI, and Category. This database will play a key role in the training process of the machine learning models used by the classification of categories, as shown in 1.

4.2 Analysis of the Database

Bibliometrix is a computational tool for bibliometric review analysis, it is one of the RStudio packages. The information regarding the database compiled and with the filters resulted in 7781 articles in the period from 2014 to 2023, with a growth rate of 12.18% of publications per year, about 16,141 authors, 29.97% with international co-authorship, 284,418 references, and 20,055 keywords.

In the period 2014 to 2023, the country that published the most was China, with 8564 articles published, followed by the USA with 2624 articles and the UK with 1289 articles. Brazil has about 729 articles published, ranking eighth among countries that have published the most on the subject among the six categories.

Regarding the main topics covered in articles, between the period 2014 to 2016, the main topics covered in the articles are industry, construction, productivity, design, optimization. From 2017 to 2019, the main topics addressed are performance, design, innovation, data envelopment analysis, productivity, and fermentation. From 2020 to 2023 the main topics covered are productivity, systems, innovation, biomass, management, and optimization. Among the categories covered in the survey, only innovation appeared as one of the main themes of the 7,781 articles. Among the two basic keywords of the research, industry, and productivity, it was found that between 2014 and 2016 industry is a more prominent theme and during the period from 2017 to 2023 productivity is the most addressed theme.

In terms of the country of the corresponding authors for the articles, China stands out with the highest number of authors, totaling 2361, followed by the USA with 674, and India with 361 authors. Additionally, China also leads in terms of the number of citations received, boasting an impressive 39853 citations, which is more than double the number of citations secured by the USA, ranking second with 15486 citations. These statistics demonstrate China's significant dominance not only in authorship and article production but also in the relevance and productivity within the industry, particularly across the six research categories. The data clearly indicates that China's research efforts in the focused theme surpass that of other countries, cementing its leading position in the field.

5 CATEGORY CLASSIFIER OF SCIENTIFIC ARTICLES

In this chapter, the process adopted to train and evaluate four machine learning models (Multinomial Naive Bayes Classification, Stochastic Gradient Descent, Support Vector Machines and Decision Tree Classifier) for classifying scientific articles in six categories will be presented: Innovation, Sustainability, Industry 4.0, Industry 5.0, Efficiency and Artificial Intelligence.

5.1 Data Preparation

In this study, we curated the Web of Science (WoS) database by excluding articles lacking essential information and ensuring that each article belonged to only one category to prevent duplication. This data cleanup process maintained the integrity of our training and validation sets, as depicted in Figure 2.

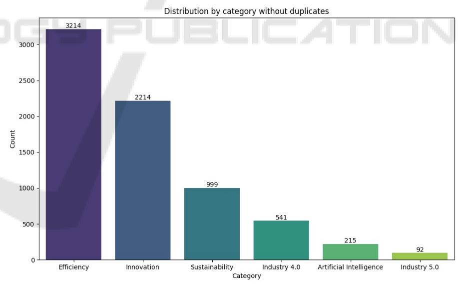

Figure 2: Distribution by category without duplicates.

The dataset was then randomly divided into 70% for training and 30% for validation, ensuring that both datasets were representative of the six categories.

For the textual representation of the articles, a corpus was built using the TF-IDF approach (Term Frequency-Inverse Document Frequency) to convert the text into numerical vectors. Furthermore, stemming was applied to reduce the words to their stem, which helped to reduce the dimensionality of the feature space and improve the performance of the models.

5.2 Machine Learning Models

With the training set prepared, the four selected machine learning models were trained using the training data.

1. Multinomial Naive Bayes Classification (MNBC) According to Chebil et al. (2023), as MNBC uses the Frequency Estimate parameter to highlight the frequencies of the available date.
2. Stochastic Gradient Descent (SGD) is observed as a classifier that learns functions of increasing complexity and generalizes over parameterized data (Nakkiran et al., 2019).
3. Support Vector Machines (SVM) Overall, SVM takes a non-linear input set and converts it to linear with the help of a kernel function (Leong et al., 2021).
4. Decision Tree Classifier (DTC) Wang et al. (2020) define DTC as a hierarchical classifier, as it provides multi-level classification and provides the specific pattern to which each data belongs, in addition to allowing flexibility against binary and multi-class classifications.

For each model, a hyper-parameter adjustment was performed using the GridSearchCV technique, which consists of defining a grid of possible values for the hyper-parameters and performing an exhaustive search for the optimal combinations.

5.3 Method Validation Metrics

Model validation is crucial in machine learning development, assessing performance on unseen data through metrics like recall, precision, accuracy, F-measure, and confusion matrices. Careful analysis of these aspects helps select the best model and fine-tune parameters for optimal performance, considering problem-specific requirements and class distinctions.

6 RELEVANCE CLASSIFIER OF SCIENTIFIC ARTICLES

Top articles were identified using citation count, and outlier detection techniques like z-scores and standard deviation were employed for selection.

7 DISCUSSION OF RESULTS

This section presents the results of the research by presenting and analyzing the performance of four different classification models applied to the task of categorizing scientific articles.

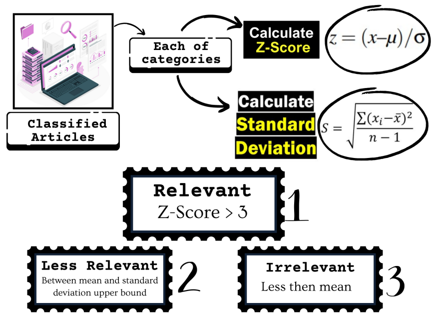

Figure 3: Relevance Classifier of Scientific Articles.

7.1 Category Classifier

The models evaluated are Multinomial Naive Bayes (MNB), SGD Classifier, Support Vector Machine (SVM), and Decision Tree. Each model was run with different hyper-parameter settings and their performance was evaluated using metrics such as accuracy, precision, recall and F-measure. The aim is to identify the best model that fits the proposed problem of subject categorization. For this analysis, data that has been pre-processed and transformed into vectors has been used. The corpus consists of documents labeled with different subject categories, and the objective is to train the models to correctly classify new documents into their respective categories.

Table 1: Web of Science results.

Variables	MultinomialNB (Naive Bayes)	SGDClassifier	SVM	Decision Tree
Accuracy	0.6748	0.7380	0.7494	0.7554
Precision	0.6542	0.7270	0.7384	0.7378
Recall	0.6748	0.7380	0.7494	0.7554
F-measure	0.6606	0.7134	0.7367	0.7441

Table 1 shows the values of the metrics for each of the models. It is also possible to see the performance for model, especially, the accuracy rate arrive 75% by Decision Tree.

1) Multinomial Naive Bayes (MNB) is a classifier based on Bayes' theorem and is widely known for its simplicity and computational efficiency. When we evaluated the performance of MNB on our subject categorization task, we found that both accuracy and recall were the same, around 67.48%. These results indicate that the model correctly classified approximately 67.48% of the instances into their correct categories. Although this is a reasonable result, we cannot consider it exceptional.

When analyzing the specific accuracy of MNB, we confirm that it is 67.48%, and the F-measure is 66.06%.

Despite its advantages in terms of simplicity and computational efficiency, the results suggest that MNB may not be the best choice for our specific task of subject categorization. In this context, it is essential to consider alternative models that have obtained more promising results, in order to select the best model to meet the requirements and objectives of our project.

2) The Stochastic Gradient Descent (SGD) classifier is a linear classifier that uses stochastic gradient descent computation for optimization. When evaluating its performance on the subject categorization task, our results show that the SGD classifier achieved an accuracy of 73.80%, a precision of 72.71% and a recall of 73.80%. In addition, the F-measure was 71.34%. These results indicate that the model outperformed MNB, which can be attributed to its ability to better handle more complex classification problems.

The recall of 73.80% indicates that the SGD classifier was able to retrieve approximately 73.80% of the categories of each item. The precision of 72.71% means that approximately 72.71% of the classified items actually belonged to that category, i.e. there was a significant proportion of true positives.

These superior results in terms of accuracy, recall and precision compared to MNB can be attributed to the SGD classifier's properties and ability to deal with linear relationships between problem features, which can be particularly useful in more complex classification tasks. Therefore, the results indicate that the SGD classifier is a more promising option for the task of subject categorization compared to MNB, but we should still consider the other models evaluated to identify the best choice that meets the specific requirements of this study.

3) Support Vector Machines (SVM) is a widely used classifier in classification problems and is known for its effectiveness in high-dimensional spaces. Our results show that the SVM achieved an accuracy of 74.94%, a precision of 73.84% and a recall of 74.94%. In addition, the F-measure was 73.67%. These results are promising and show that SVM outperformed both MNB and SGD Classifiers.

The recall of 74.94% indicates that the SVM was able to recover approximately 74.94% of the classified categories. In turn, the precision of 73.84% shows the percentage of classified categories that actually belonged to that category, demonstrating a significant proportion of true positives.

These superior results in terms of accuracy, recall, and precision compared to MNB and SGD Classifier confirm the effectiveness of SVM for the task of subject categorization, especially in scenarios with high data dimensionality and classification complexity.

Therefore, the results highlight SVM as a highly competitive and successful option in the subject categorization task. However, it is still crucial to consider other factors, such as runtime and interpretability, in order to select the most suitable model for the specific needs and requirements of our project.

4) Decision Tree is a decision rule-based model known for its interpretability. Our results show that the Decision Tree achieved an accuracy of 75.54%, a precision of 73.78%, and a recall of 75.54%. In addition, the F-measure was 74.41%. These results are the best of the models tested, indicating that Decision Tree excelled at the task of subject categorization.

The recall of 75.54% means that Decision Tree was able to correctly retrieve most of the categorized categories. The precision of 73.78% means the percentage of classified categories that actually belonged to that category, which is a positive result.

These satisfactory results in terms of accuracy, recall, and precision confirm that the Decision Tree is a competitive option for the task of subject categorization. Moreover, its interpretability is a significant advantage, allowing a better understanding of the decisions made by the model and facilitating the identification of classification patterns.

Based on the results presented, we can conclude that Decision Tree was the model that obtained the best performance for the subject categorization task. Its accuracy of 75.54% and F-measure of 74.41% outperformed the other models evaluated. The interpretability of Decision Tree can be a significant advantage when dealing with subject categorization.

7.2 Relevance Classifier

After classifying the 17,000 articles into subject categories, they were further grouped into three categories of relevance: Irrelevant, Less Relevant, and Relevant. The categorization was based on their citation counts using statistical metrics such as the mean and standard deviation and Z-score.

Articles with citation counts lower than the mean were classified as "Irrelevant." Articles with citation counts falling between the mean and the upper bound of the standard deviation were labeled as "Less Relevant." On the other hand, articles with a z-score greater than 3, indicating a substantial deviation from the mean, were classified as "Relevant." This approach allowed for a comprehensive and nuanced assessment of the articles' significance and impact within their respective subject categories.

Below is a table 2 displaying the distribution of articles across each category:

In the table 3 in the appendix, we present a table

Table 2: Quantity of articles per category.

Category	Irrelevant	Less Relevant	Relevant
Innovation	6317	473	253
Efficiency	3606	353	238
Sustainability	1181	97	59
Industry 4.0	1232	98	53
Artificial Intelligence	184	10	7

showcasing the names of the top 5 most relevant articles within each category:

8 CONCLUSION

In conclusion, this paper aimed to compare systematic literature review methods to identify the most efficient approach for conducting quick and effective literature reviews. We categorized productivity in the industry into six principal sectors, namely Innovation, Sustainability, Industry 4.0, Industry 5.0, Efficiency, and Artificial Intelligence. Subsequently, we conducted systematic literature reviews for each of these sectors.

To achieve efficient categorization, we employed four Machine Learning models: Multinomial Naive Bayes Classification (MNBC), Stochastic Gradient Descent (SGD), Support Vector Machines (SVM), and Decision Tree Classifier (DTC). Among these models, the Decision Tree Classifier (DTC) demonstrated superior performance with a classification accuracy exceeding 75%.

To identify the most relevant contributions of scientific articles in each research domain or category, we utilized a relevance classifier with two essential metrics: the standard deviation of the mean and the Z score.

As a forward-looking suggestion for further research, we highlight the significance of incorporating additional data sources and considering the h-index as an alternative approach for ranking the relevance of articles. By expanding the scope of data and incorporating diverse evaluation metrics, future studies can enhance the accuracy and depth of systematic literature reviews, ultimately contributing to more comprehensive insights and informed decision-making in the domain of industrial productivity.

ACKNOWLEDGMENTS

Thanks to the University of Brasilia (UnB) for supporting this research and to the National Industrial Learning Service (SENAI) of the National Confederation of Industry (CNI) for partially supporting this research.

REFERENCES

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of educational research*, 80(2):207–245.

Borracini, G. F. (2021). O impacto dos subsídios na evolução da produtividade do setor agrícola: e obtenção do título de mestre no mestrado profissional em economia. *IDP/FAB*.

Chebil, W., Wedyan, M., Alazab, M., Alturki, R., and Elshaweesh, O. (2023). Improving semantic information retrieval using multinomial naive bayes classifier and bayesian networks. *Information*, 14(5):272.

Cohen, A. M., Hersh, W. R., Peterson, K., and Yen, P.-Y. (2006). Reducing workload in systematic review preparation using automated citation classification. *Journal of the American Medical Informatics Association*, 13(2):206–219.

de Melo, M. K., Faria, A. V. A., Weigang, L., Nery, A. G., de Oliveira, F. A. R., Barreiro, I. T., and Celestino, V. R. R. (2022). Few-shot approach for systematic literature review classifications. *Proceedings of the 18th International Conference on Web Information Systems and Technologies WEBIST*.

Galvão, M. C. B., Pluye, P., and Ricarte, I. L. M. (2017). Métodos de pesquisa mistos e revisões de literatura mistas: conceitos, construção e critérios de avaliação. *InCID: Revista de Ciência da Informação e Documentação*, 8(2):4–24.

Howick, J. (2011). The oxford 2011 levels of evidence. *http://www. cebm. net/index. aspx? o= 5653*.

Leong, W. C., Bahadori, A., Zhang, J., and Ahmad, Z. (2021). Prediction of water quality index (wqi) using support vector machine (svm) and least square-support vector machine (ls-svm). *International Journal of River Basin Management*, 19(2):149–156.

Mundial, B. (2018). *Emprego e crescimento: a agenda da produtividade*. Brasília: Grupo Banco Mundial.

Nakkiran, P., Kaplun, G., Kalimeris, D., Yang, T., Edelman, B. L., Zhang, F., and Barak, B. (2019). Sgd on neural networks learns functions of increasing complexity. *arXiv preprint arXiv:1905.11604*.

Santos, M. R. d., Pandullo, R. P., and Souza, E. C. d. (2019). Evolução da produtividade total dos fatores e da produtividade da mão de obra da indústria brasileira: uma análise setorial e comparativa. *Universidade Católica de Brasília*.

Torezani, T. A. (2018). Evolução da produtividade brasileira: mudança estrutural e dinâmica tecnológica em uma abordagem multissetorial. *Universidade Federal do Rio Grande do Sul*.

Wang, F., Wang, Q., Nie, F., Li, Z., Yu, W., and Ren, F. (2020). A linear multivariate binary decision tree classifier based on k-means splitting. *Pattern Recognition*, 107:107521.

Weigang, L., Enamoto, L. M., Li, D. L., and Rocha Filho, G. P. (2022). New directions for artificial intelligence: human, machine, biological, and quantum intelligence. *Frontiers of Information Technology & Electronic Engineering*, 23(6):984–990.

APPENDIX

Table 3: Relevance articles extracted from WoS classified by category.

Category	Title
Artificial Intelligence	Brave new world: service robots in the frontline
	Industry 5.0-A Human-Centric Solution
	Agriculture 4.0: Broadening Responsible Innovation in an Era of Smart Farming
	Industrial internet of things: Recent advances, enabling technologies and open challenges
	From Industry 4.0 to Agriculture 4.0: Current Status, Enabling Technologies, and Research Challenges
Efficiency	Impact of energy conservation policies on the green productivity in China's manufacturing sector: Evidence from a three-stage DEA model
	The Role and Impact of Industry 4.0 and the Internet of Things on the Business Strategy of the Value Chain-The Case of Hungary
	Energy and CO2 emissions performance in China's regional economies: Do market-oriented reforms matter?
	Modeling the role of environmental regulations in regional green economy efficiency of China: Empirical evidence from super efficiency DEA-Tobit model
	Enhancing microalgal biomass productivity by engineering a microalgal-bacterial community
Industry 4.0	State-of-the-art in surface integrity in machining of nickel-based super alloys
	The cost of additive manufacturing: machine productivity, economies of scale and technology-push
	The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda
	Multi-response optimization of minimum quantity lubrication parameters using Taguchi-based grey relational analysis in turning of difficult-to-cut alloy Haynes 25
	A Survey on Industrial Internet of Things: A Cyber-Physical Systems Perspective
Innovation	Innovation in the pharmaceutical industry: New estimates of R&D costs
	Influence of tribology on global energy consumption, costs and emissions
	Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry
	Environmental regulation and competitiveness: Empirical evidence on the Porter Hypothesis from European manufacturing sectors
	Different Types of Environmental Regulations and Heterogeneous Influence on Green Productivity: Evidence from China
Sustainability	Natural fiber reinforced polymer composites in industrial applications: feasibility of date palm fibers for sustainable automotive industry
	Nanotechnology in Sustainable Agriculture: Recent Developments, Challenges, and Perspectives
	'Green' productivity growth in China's industrial economy
	Sustainable manufacturing in Industry 4.0: an emerging research agenda
	Can carbon emission trading scheme achieve energy conservation and emission reduction? Evidence from the industrial sector in China