Machine Learning Application in Battery Prediction: A Systematic Literature Review and Bibliometric Study

Azure Kamul[1*], *Nico* Hananda[1], *Christian* Harito[1], *Endang* Djuana[2], *Gregorius Natanael* Elwirehardja[3,4],
Bens Pardamean[3], *Fergyanto E.* Gunawan[5], *Arief S.* Budiman[5], *Muhammad* Asrol[5], *A. A. N. Perwira* Redi[5], *Tim* Pasang[6]

[1]Industrial Engineering Department, Faculty of Engineering, Bina Nusantara University, Jakarta, 11480, Indonesia
[2]Electrical Engineering Department, Faculty of Industrial Technology, Universitas Trisakti, Jakarta, 11440, Indonesia
[3] Bioinformatics and Data Science Research Center, Bina Nusantara University, Jakarta, Indonesia, 11480
[4] Computer Science Department, BINUS Graduate Program - Master of Computer Science Program, Bina Nusantara University, Jakarta, Indonesia, 11480
[5]Engineering Industrial Engineering Department, BINUS Graduate Program - Master of Industrial Engineering, Bina Nusantara University, Jakarta, Indonesia 11480
[6]Department of Mechanical and Manufacturing Engineering and Technology, Oregon Institute of Technology, Klamath Falls, OR 97601, USA

Abstract. Recently, the popularity of li-ion batteries has attracted many researchers to carry out the battery's maximum potential. Predicting batteries condition and behavior is part of the process that is considered challenging. ML algorithm is widely applied to overcome this challenge as it demonstrates a successful outcome in optimizing the complexity, accuracy, reliability, and efficiency of battery prediction. Yet, we believe there is a particular research area of battery prediction that can further be explored and enhanced with machine learning capability. Therefore, we perform a systematic literature review and bibliometric study to uncover the gap in the machine learning application in the battery prediction field. This study is divided into four stages: (1) literature search from the Scopus Database, (2) filtering the results based on keywords and prepared criteria using PRISMA method, (3) systematic review from filtered papers to provide further understanding, and (4) bibliometric analysis from visualization created in VOSViewer *software*. The analysis findings determine battery safety and performance prediction as a potential gap in the scope of machine learning for battery prediction research and provide some insightful information to assist future researchers. We envision this study to encourage further battery research, which will assist in the creation of better, cleaner, safer, and long-lasting energy resources.

1. Introduction

The energy demand has significantly grown in the past few decades. However, carbon-based energy resources that are massively used severely impact the ecosystem [1]. Data shows that global energy consumption has resulted in a 45% increase in CO_2 emissions from 2000 to 2019. Renewable and low-emissions energy resources are urgently in the call to mitigate CO_2 emissions by replacing the currently used carbon-based energy resources, such as fossil fuels [2]. Batteries, especially Li-ion types, without a doubt, are known as one of today's most favorable alternatives due to their efficiency and flexibility. The popularity of the Li-ion battery is recognizable by its massive market size reaching 41.97 billion dollars in 2021. A compound annual growth rate of 18.1% is anticipated to enlarge the Li-ion battery market by 2030 [3]. With the increasing attention to batteries, the expectation of cost, safety, lifetime, performance, reusability, and recyclability improvement has yet to stop.

Enormous battery R&D has been conducted to boost the progression of improvement. This is further affected by the successful development of electric vehicles (EVs), which highly depends on Li-ion batteries. Specifically, the EVs industry desperately requires precise battery lifetime prediction to estimate warranty costs for electric vehicles and grid storage applications which could reduce battery deployments cost [4]. Other predictions, including cell performance, safety, aging, and battery health, are significant concerns for battery application in specific domains. Researchers should evaluate several battery state parameters with a particular method to ensure the reliability of battery prediction. Among those, remaining useful life (RUL), battery state of health (SOH), and state of charge (SOC) [5] are considered the main parameter in battery management systems which able to enhance the operation of batteries (Toughzaoui et al., 2022). The result from the parameter estimation can be used as beneficial

---

* Corresponding author: azure.kamul@binus.ac.id

information to predict when the battery should be removed or replaced.

Battery research required considerable efforts to achieve reliable predictions until technological advancement has influenced battery research. The present research is mainly supported by technology to develop exact models that allow early preventive alert, dependable interpretation, and broader application for cycling conditions. Machine learning (ML) as a branch of artificial intelligence (AI) is part of those advancements that are widely adopted and employed nowadays in battery R&D. ML has large-scale capabilities to compute multivariable data set, discover a pattern, and unlock application that is hardly determined by other methods [4]. Thus, AI and ML bring a new era of data-driven predictive analysis approaches by efficiently overcoming the challenges of battery research, which usually deals with an immense number of variables and data. Furthermore, enough high-quality data can be used to develop a well-trained ML algorithm capable of simulating large-scale experimental data with a high level of accuracy, making it valuable for battery prediction.

This paper is prepared to uncover the research and publication gap on machine learning for battery prediction. To do so, the researchers perform a bibliometric analysis and systematic literature review relying on the database of Scopus-indexed publications. The researchers also implement Publish and Perish and VOSViewer applications to support the process. The aim is to identify and evaluate methods or algorithms of machine learning being used to predict issues on particular types of batteries and provide recommendations on which field could be more beneficial in future research on battery prediction.

2. Literature Review

2.1. Battery diagnosis and prognosis,

Li-ion batteries have gained more interest in the battery industry thanks to their high energy density, cost-effectiveness, and long-lasting life. However, even with its extraordinary advantages, there are currently no possibilities for Li-ion batteries to avoid degradation. Degradation refers to a gradual decline in battery performance, such as self-discharge, disproportion, and loss in cell capacity [7]. Due to aging, environmental effects, and dynamic loading over battery lifetime, degradation severely limits battery functionality. This is where battery diagnosis comes in handy. The task is to track the underlying degradation and pull out countermeasures to impede and prevent any developing fault, which sooner can result in life-threatening issues such as explosions due to overheating and short-circuiting. Meanwhile, battery prognostics deals with predicting remaining useful life (RUL) to estimate how much longer a battery will reach its ground-level performance starting when the degradation has been detected [8]. RUL is described as the battery's remaining load cycles before reaching its end of life (EoL) [9]. The

prediction is essential as the battery should be replaced at a certain point after degradation to ensure the user's safety. Using a battery after reaching its end of life will lead to some inevitable severe battery failure which battery diagnoses are craved to prevent. By accurately predicting the RUL of a battery, maximum life expectancy can be estimated, which leads users to utilize the battery's fullest potential until it reaches its furthest dying condition. Besides RUL, other battery parameters including state of energy (SOE), state of power (SOP), state of health (SOH), state of function (SOF), state of charge (SOC), state of balance (SOB), state of temperature (SOT), and remaining discharge time (RDT) are also considered when predicting battery behavior [5]. Among those, SOH and SOC are considered the main parameters in battery management systems that can optimize battery operation (Toughzaoui et al., 2022). SOC defines the remaining charge percentage of a battery compared to its fully charged state capacity. The information gained from accurate SOC estimation is beneficial in optimizing battery operation strategies and cell balancement in a battery pack. Along with battery resistance, SOC is widely used to calculate the SOH of Li-ion batteries [4]. SOH, instead, describes the fully charged stated capacity of a battery compared to its capacity in brand new condition after being manufactured. Prediction of SOH can be implemented in a battery management system for online monitoring, where users can track the battery performance and schedule any repairs or replacements beforehand [11]. The difference between those two parameters is SOC can reach 0% condition, while SOH does not. Practically, SOC is 100% when a battery is charged to its entire state and 0% when the battery charge is used to empty. On the other hand, the battery has 100% SOH when freshly manufactured and reaches 80% at the end of life [12].

Unfortunately, it is a back-breaking job to accurately predict battery parameters due to uncertain environmental effects and conditions. Additionally, the aforementioned parameters are interval variables that are difficult to estimate even with a sensor. Thus, constructing a degradation model to estimate battery state parameters accurately is crucial in battery prediction [13]. Battery prediction can be in terms of health, aging, safety, and performance [4]. State prediction from an accurate model can ensure operation reliability, battery system optimization, and safety management substance for a battery. In general, battery state estimation is split into three different approaches: direct, model-based, and data-driven method. The direct measurement method commonly deals with direct measurement and look-up table approach such as internal resistance, open-circuit voltage (OCV), etc. The model-based method is divided into a filter-based method such as particle filter (PF) and Kalman filter (KF) as well as observer-based methods such as sliding mode, Luenberger, and H-infinity. Lastly, the data-driven method, which specifically utilizes technology, alternately called the machine learning method, consists of fuzzy logic, neural network, support vector machine, genetic algorithm, etc. Further explanation of each method can be found in [14]. The machine learning method's capability to translate high-

dimensional and noisy environmental data into understandable information for battery diagnostic and prognostic makes it today's most powerful and convenient method for battery prediction [15].

2.2. Machine learning in battery R&D

Battery R&D is a complex multivariable problem dealing with an enormous amount of data. The trial-and-error approach and interpretation of those data are widely being made in present battery R&D. Still, doing that level of trial and interpretation is up to a point where the human brain, time, and energy are not capable. This is when researchers need technology assistance, particularly machine learning, to efficiently optimize the Battery R&D process [4]. Machine learning is a category of Artificial Intelligence that can learn and understand a structure of data and design that particular data into a model that people can understand and utilize to make certain decisions. Machine learning learns and improves data or experience through a particular method and algorithm, which eventually provides specific solutions without involving us in the process [16]. The accuracy and reliability of machine learning algorithms are highly dependent on the input data's quantity, quality, and integrity.

There are four major different machine learning methods: supervised, unsupervised, semisupervised, and reinforcement methods. Supervised ML is used to train a machine to understand the relationship between input and output data by learning a labeled data set. The aim is to employ the data from the training process to develop a numerical model linking some input to certain outputs. The machine will later learn that, given certain information, it will predict an appropriate outcome based on the model created [17]. On the other hand, unsupervised ML learns from data set containing input data without corresponding output to learn a pattern without any specific feedback. The purpose is to identify data groups or useful variables from the data or so-called clustering [17]. The main difference between these two methods is unsupervised ML does not have any historical information about the input and output relationship of the data, while supervised does. Thus, it is up to the operator who uses the unsupervised method to assess whether the result is perspectively true or false [18]. The semisupervised method combines supervised and unsupervised, meaning ML will learn through data sets with labeled and unlabeled data. More accurate capture of various synthesis procedures will be featured in the machine learning algorithm with this combination method. Thus, humans can easily interpret and understand the presented result [19]. At last, reinforcement is a method that enables machine to learn based on reward and punishment. Training will be done through trial-and-error that rewards the machine for the desired result and punishes it for undesired results. The purpose is to develop a machine that is automatically capable of observing its environment based on a particular condition. This type of method is typically beneficial in the field of robotics, self-driving task, and other automation processes

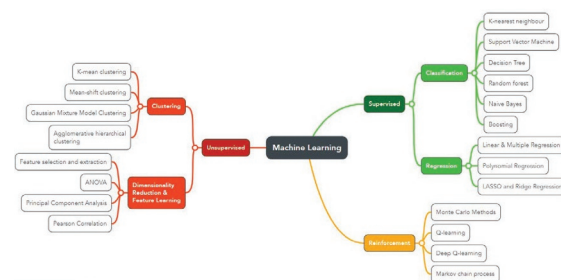[20]. Several algorithms from each method are shown in Figure 1.

Fig.1. Machine Learning Method & Algorithm Mind Map [20]

Selecting the most appropriate machine learning method and algorithm is challenging, especially in the battery R&D. Various factors must be considered, including the data availability, desired result, and required model. But machine learning has come up as a promising tool for battery prediction by estimating battery state, including state of charge, state of health, and remaining useful life of batteries. A neural network is probably the dominant algorithm for SOC estimation due to its accurate result prediction. However, preferred machine learning algorithms for either SOC, SOH, or RUL are still varied [12].

2.3. Bibliometric analysis and software

The interest in the bibliometric analysis as a quantitative method for scientific publication has been proven as researchers continuously adopt it. In Scopus only, the results from searching bibliometric analysis terms (14-5-2022) show that implementation of this method in 2022 has reached 1.238 publications, almost twice the number of 2017 results. Moreover, the search is being done in May, which means the result in 2022 might increase even more until the end of the year. Bibliometric analysis is a quantitative method that allows researchers to utilize enormous bibliometric data (e.g., publication and citation) to measure and learn something new about scientific research. It is best known for the purpose of identifying and summarizing emerging trends in a particular research field [21]. There are four steps needed to execute a proper bibliometric analysis. First, the researcher must define the purpose and scope of the bibliometric study. The scope of the bibliometric study is expected to be broad as bibliometric analysis is designed to reveal patterns from massive data. Thus, the researcher needs to review the number of related papers available. If the available number is less than 100, conducting bibliometric analysis is not a recommended action. Second, the researcher needs to select proper techniques for bibliometric analysis according to the scope of the study. Third, the researcher must collect an appropriate amount of data. In this research, data are collected from the Scopus database with the support of Publish or Perish applications. Finally, the researcher must perform a

bibliometric analysis and conclude the result. The result can either be performance analysis that summarizes the contribution from collected research data to the aim of study or science mapping that summarizes the relationship from collected research data to the of study [22]. Several analytical tools such as VOSViewer have been developed to simplify the process by enabling bibliometric data mapping through visualization and network approaches. With bibliometric data mapping, analysis can further uncover the relationship among scholarly research by manipulating several aspects, including network approach, size, nodes, and interaction. This is because of the bibliometric mapping capability to quantify details such as cluster, direction, and topic of a certain field's knowledge [23].

3. Methods

The researcher accessed the Scopus database (https://www.scopus.com/) on 18 May 2022 to gather desired literature data to perform the systematic literature review and bibliometric analysis of this study. The selection of Scopus as the source to retrieve relevant articles for this study is because of its well-known indexed database containing massive numbers of citations, abstracts, articles, journals, conference papers, and books. Several keywords in the scope of machine learning for battery prediction were carefully selected to gather desired and relevant literature data. The keyword used in performing literature search includes "Machine Learning", "Battery", "Prediction", and "Algorithm". Systematically, the workflow of literature search in this study is done with PRISMA method as shown in Figure 2.
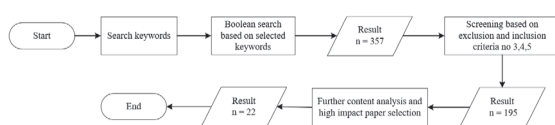

Fig.2. Workflow of Literature Search

Literature searches based on selected keywords were done in the form of a Boolean search in the Scopus search engine written as "TITLE-ABS-KEY (machine AND learning AND battery AND prediction AND algorithm)". After inputting selected keywords in Boolean search, the researcher found over 357 publication results from Scopus databases. The result from searching selected keywords was further reduced through a set of inclusion and exclusion criteria shown in Table 1. With the help of the "Limit To" syntax in the Scopus search engine, researchers could limit the results based on the third, fourth, and fifth inclusion and exclusion criteria. Additional syntax was added to the Boolean search written as "TITLE-ABS-KEY (machine AND learning AND battery AND prediction AND algorithm) AND PUBYEAR > 2014 AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (LANGUAGE,"English"))". The added syntax from those three criteria reduced the results from 357 to 195 publications. Finally, Publish or Perish

search and content analysis based on the first and second criteria was performed to discover high impact papers with great relevancy to the scope of machine learning for battery prediction. This step narrowed down the results into 22 papers, where 19 of them are acknowledged as key papers. The researcher then applied bibliometric analysis to the 22 selected papers using the VOSViewer application to visualize the network and clustering of several keywords obtained in the papers. In addition, a systematic literature review of the 22 papers is also carried out to give more in-depth knowledge about machine learning for battery prediction.

Table 1. Literature Search Inclusion/Exclusion Criteria

No	Inclusion	Exclusion
1	The scope of the research must focus on battery prediction	Other energy resources study were excluded
2	Only research adopting machine learning method/algorithm were selected	Research adopting other techniques beside machine learning were excluded
3	Results must be in article type	Other type of results were excluded
4	Articles selected must be written in english language	Other language written articles were excluded
5	Articles selected must be published between 2015-2022	Articles published older than that year were excluded

4. Results and Discussion

4.1. Systematic Literature Review

From the 22 selected Scopus-indexed articles related to machine learning for battery prediction, researchers summarized them based on the used algorithm, parameters, and results shown in Table 2. There are various types of machine learning algorithms used in the selected articles, such as Extreme Learning Machine (ELM), Neural Network, Support Vector Machine, Relevant Vector Machine (RVM), Gaussian Process, Random Forest, k-Nearest Neighbor, Kalman Filter, Bayesian Ridge Regression, Kernel Function, Decision Tree, etc. Still, almost all the articles mainly highlighted an improved version of the algorithm by combining it with a certain method and approach. For instance, the proposed RVM algorithms in the selected papers are enhanced with incremental learning [24], Mixed Kernel Function [25], Multiple Kernel Function [26], Selective Kernel Function [27], Genetic Algorithm [28], and Kalman Filter [29] to produce a more accurate and reliable prediction. An improved version of the ELM algorithm is also shown in several articles. [30], for example, proposed an improved ELM algorithm called adaptive online sequential extreme learning machine, which is proven to have better consistency and accuracy in SOC prediction as well as reasonable training time and required input data. Both RVM and ELM are recognized as the most dominant algorithms applied in the 22 selected papers. Additionally, conventional machine learning algorithms are typically used to do a comparative study, which later spotlights the proposed improved method advantages mentioned in the article. In terms of parameters, most of the selected articles take SOC, SOH, and RUL as the main parameter in their prediction aim. Other parameters include cathode material [31] and temperature [32]. To sum it up, the primary focus of battery prediction is

generally on its lifetime, health, and capacity. Some battery prediction research is also implemented in real-life components such as the lifetime of IoT networks battery [33] and EVs driving range [34].

Table 2. The summary of 22 selected articles regarding machine learning algorithm in battery prediction

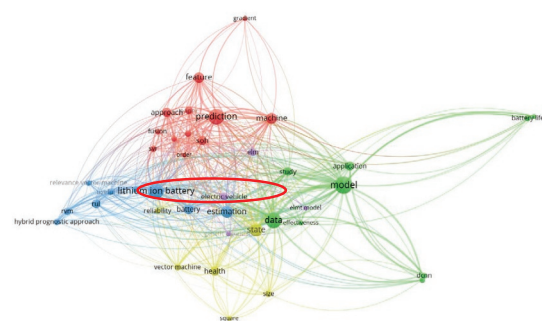| Reference	Type(s) of Algorithm	Parameters	Results
(Chin & Gao, 2018)	AOS-ELM (Adaptive Online Sequential-Extreme Learning Machine)	SOC	AOS-ELM could consistently make SOC predictions w smaller root mean squared error (RMSE) in reasonal training time than other ELM approaches. AOS-ELM a compatible with limited and sequential data
(Hong et al., 2020)	Long-Short Term Memory (LSTM) Neural Networks	SOC	LSTM can perform accurate and fast multi-forward-st SOC, which eliminates driving anxiety by diagnosing a preventing SOC anomalies in EVs batteries beforehanc
(Hu et al., 2021)	Artificial Neural Network (ANN), Support Vector Machine (SVM), Relevant Vector Machine (RVM), & Gaussian Process Regression (GPR)	SOH	Comparison of 4 ML algorithms accompanied w different feature extraction and reduction are applied to battery health predictions based on SOH estimation. T results show that the combination of GPR and fusic based selection method have the best accuracy a computational efficiency.
(Khumprom & Yodo, 2019)	Deep Neural Network, K-Nearest Neighbor, SVM, ANN, Linear Regression	RUL & SOH	The result shows that DNN outperforms all otr algorithms in SOH and RUL prediction accuracy. T accurate prognostic result is crucial to know the best tir to replace the battery before it causes severe failure. T downside of DNN is its computational time, which mig not be suitable for real-time processing.
(D. Liu et al., 2015)	Incremental Optimized Relevance Vector Machine (IP-RVM)	RUL	The proposed IP-RVM algorithm has a better RI prediction precision even when compared to a re-train off-line RVM algorithm. Its operating time is also reduc almost by half. Yet, IP-RVM needs to be further evaluat to reduce the uncertainty of the RUL estimation.
(C. Ma et al., 2019)	Multiple-View Feature Fusion Support Vector Regression Ensemble Strategy (MVFF-ESVR)	SOH	MVFF-ESVR fuses multiple trained SVR by utiliziı AdaBoost and Stacking algorithms to generate two SC predictors. The result shows extremely low err compared to other existing predictors mainly because its extracting and fusing features capabilities that prev information loss in traditional extraction features (TFE
(Y. Ma et al., 2020)	Broad Learning-Extreme Learning Machine (BL-ELM)	SOC & Cycle life	BL-ELM is mapping input data to produce feature noc that are further enhanced without increasing the numl of layers. This increases the prediction ability wh avoiding growth in required input data. Their esult sho the model can predict capacity and cycle life mc accurate, with an average time of 1,48s
(Maddikunta et al., 2020)	PCA-Based Random Forest Regression	Battery life	The result from PCA-based Random Forest regression predicting IoT's network battery proves the accuracy a superiority of the proposed model to other regressi algorithms, such as linear and XGBoost regression.
(Mawonou et al., 2021)	Random Forest Approach	SOH	The proposed algorithm can predict SOH with estimation error of 1,27%. Accurate SOH estimation useful in automotive, specifically EVs, to predict steady range and available power over lifetime.
(Roman et al., 2021)	Random Forest (RF), DNN, Bayesian Ridge Regression (BRR) And GPR	SOH	Assessment of RF, DNN, BRR, and GPR for BHUN indicates that the base algorithm with lowest err depends on cell charging protocol. BHUM is consider a reliable technique that could be applied for componer requiring real-time estimation for SOH.
(Attarian Shandiz & Gauvin, 2016)	LDA (Linear Discriminant Analysis), ANN, SVM, K-NN, Random Forest & Extremely Randomized Trees	Crystal system of cathode materials	RF and extremely randomized trees have the high average accuracy compared to other ML classificati algorithms in predicting the parameters. The volume crystals and number of sites have the strongest correlati in determining the type of crystal system

4.2. Bibliometric Analysis

Figure 3. Bibliographic Overlay Visualization

Figure 4. Bibliographic Density Visualization

Fig.5. Bibliographic Network Visualization

Through VOSViewer software, a bibliographic mapping was conducted to visualize the overlay, density, and network of the keywords found in the 22 selected papers related to machine learning for battery prediction. The output is shown in Figure 3, Figure 4, and Figure 5. There are 5 clusters identified in the bibliometric network along with its main keyword, namely prediction (red) with other 11 keywords, lithium-ion battery (blue) with other 8 keywords, model (green) with 8 other keywords, state (yellow) with other 5 keywords, and electric vehicle (purple) with other 4 keywords. The blue cluster indicates that lithium-ion batteries are the most popular battery types used for research battery prediction. It is highly related to the emerging trend of electric vehicles (purple), as shown in the overlay visualization. The amount of EVs research illustrated in Figure 4 seems quite phenomenal, considering most EVs research started in 2019.5 (Figure 3). It is also visible that the red cluster is the most dominant cluster reflected from its number of keywords and network. Among 12 keywords in the red cluster, it contains the keywords "Prediction" and "Machine Learning" which is the focus of this study. This finding shows a high correlation between machine learning and prediction in battery research. Researchers hovered over the particular keyword of "machine learning" to see a more specialized connection, revealing all correlated and uncorrelated keywords, as seen in Figure 6. In terms of ML algorithm, there are two keywords correlated to machine learning, including "Vector Machine" and "ELM". The correlation of those two keywords reinforces the statement in the systematic literature review regarding RVM and ELM as foremost ML algorithms applied in battery prediction. Finally, Figure 6 provides insightful information about the research gap in machine learning for battery prediction. The visible relation of keywords "health" and "useful life" to "machine learning" points out that research on machine learning for battery prediction predominantly aims at aging and health prediction. "RUL" and "SOH" keywords as battery parameters for aging and health prediction displayed in the network further prove the previous statement. However, battery development should also accurately predict the performance and safety of batteries. Yet, the bibliometric result does not present any keywords regarding performance and safety. Therefore, it can be assumed that battery safety and performance prediction is the publication and research gap obtained from this systematic literature review and bibliometric study. Still,

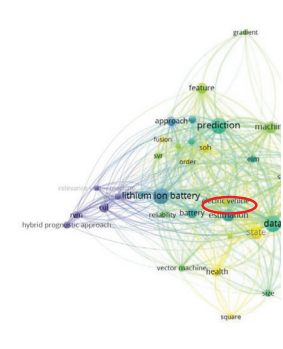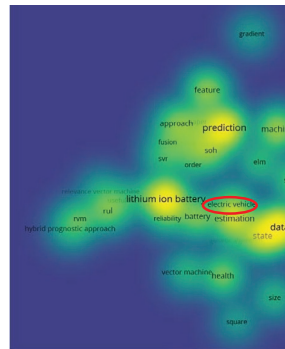to clarify, this bibliometric study only extracted keywords from the title and abstract of the 22 selected papers. The selected papers were also filtered exclusively from the Scopus database. This can be considered a limitation of the research gap obtained in this study.
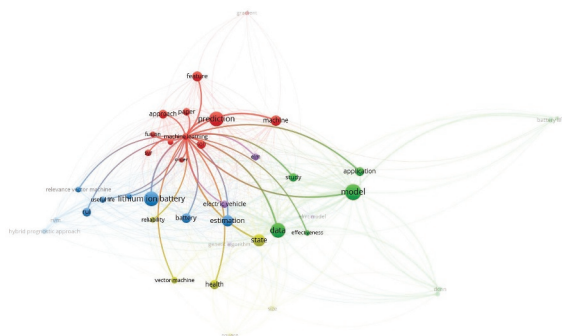

Fig.6 Bibliographic Network Visualization specialized in "Machine Learning" keywords

4.3. Future Recommendations

In this section, the researcher presents recommendations based on the systematic literature review and bibliometric results. With the increasing attention on batteries, specifically li-ion batteries, manufacturers are expected to develop more superior batteries concerning their performance, health, lifetime, and safety all at once. Machine learning as a tool to accelerate the trial-and-error process, accurately predicting all those aspects' results becomes more and more popular. Nevertheless, obtained results from this study indicate a gap in terms of machine learning for battery safety and performance prediction. Further exploration and research about battery safety and performance prediction using machine learning algorithms are recommended to close the particular gap. Few key papers related to the mentioned topic can be found in [35] and [36]. Future researchers interested in this topic might use those key papers as early references to start the study. For the suitable applied algorithm, there are numerous possibilities as many new improved and combined versions of ML algorithms are endlessly proposed. Some ML algorithms dominantly used in battery prediction are improved versions of RVM and ELM. Still, future researchers should observe the right algorithm based on the prediction aim, parameters, available data, and time. Last, we envision that future research on performance and safety prediction in the field of batteries could assist the creation of better, cleaner, safer, and long-lasting energy resources.

5. Conclusion

This paper proposes a systematic literature review and bibliometric analysis based on several papers. The analyzed paper was gained through a literature search from Scopus Database with the keywords "Machine Learning", "Algorithm", Battery", and "Prediction". Some inclusion and exclusion criteria are also taken into consideration. Through a systematic search process, researchers identified 22 core papers that possess high relevance to the keywords and fulfill the required criteria. The 22 selected papers were then summarized and discussed in the form of a systematic literature review to provide a brief understanding to the readers. In addition, bibliometric analysis is conducted to uncover the research and publication gap on machine learning for battery prediction. The results reveal machine learning for performance and safety prediction as the particular gap. There is also useful information regarding the predominantly utilized algorithms and batteries in battery prediction research. In terms of battery types, li-ion batteries are clearly gaining the most attention from researchers. On the other hand, the most utilized algorithms based on the bibliometric results are relevant vector machine and extreme learning machine, especially the improved types. As a way forward, researchers encourage further exploration and research, specifically about battery safety and performance prediction using machine learning algorithms, by providing a few papers as an early references. We believe this study could allow future researchers to carry out more potential battery prediction results by utilizing machine learning, which hopefully supports the development of the battery industry.

References

1. P. Greim, A. A. Solomon, and C. Breyer, Nature Communications 2020 11:1 **11**, 1 (2020)
2. C. Ling, Npj Computational Materials 2022 8:1 **8**, 1 (2022)
3. Grand View Research, Grand View Research 150 (2022)
4. T. Lombardo, M. Duquesnoy, H. El-Bouysidy, F. Årén, A. Gallo-Bueno, P. B. Jørgensen, A. Bhowmik, A. Demortière, E. Ayerbe, F. Alcaide, M. Reynaud, J. Carrasco, A. Grimaud, C. Zhang, T. Vegge, P. Johansson, and A. A. Franco, Chemical Reviews (2021)
5. L. Fang, J. Li, and B. Peng, Energy Procedia **158**, 3008 (2019)
6. Y. Toughzaoui, S. B. Toosi, H. Chaoui, H. Louahlia, R. Petrone, S. Le Masson, and H. Gualous, Journal of Energy Storage **51**, 104520 (2022)
7. A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, Journal of Power Sources **239**, 680 (2013)
8. J. Weddington, G. Niu, R. Chen, W. Yan, and B. Zhang, Neurocomputing **458**, 440 (2021)
9. Y. Lei, Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery 281 (2017)

10. Y. Toughzaoui, S. B. Toosi, H. Chaoui, H. Louahlia, R. Petrone, S. Le Masson, and H. Gualous, Journal of Energy Storage **51**, 104520 (2022)

11. Y. Li, K. Liu, A. M. Foley, A. Zülke, M. Berecibar, E. Nanini-Maury, J. Van Mierlo, and H. E. Hoster, Renewable and Sustainable Energy Reviews **113**, 109254 (2019)

12. M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit, and Z. W. Seh, Nature Machine Intelligence 2020 2:3 **2**, 161 (2020)

13. X. Hu, L. Xu, X. Lin, and M. Pecht, Joule **4**, 310 (2020)

14. W. Liu, T. Placke, and K. T. Chau, Energy Reports **8**, 4058 (2022)

15. A. Nuhic, T. Terzimehic, T. Soczka-Guth, M. Buchholz, and K. Dietmayer, Journal of Power Sources **239**, 680 (2013)

16. B. H and R. H, International Journal of Engineering Applied Sciences and Technology **6**, (2021)

17. S. Russell and R. Norvig, *Artificial Intelligence: A Modern Approach, 4th Edition*, 4th ed. (Pearson Education, 2020)

18. M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, 3 (2020)

19. H. Huo, Z. Rong, O. Kononova, W. Sun, T. Botari, T. He, V. Tshitoyan, and G. Ceder, Npj Computational Materials **5**, (2019)

20. I. H. Sarker, Sn Computer Science **2**, (2021)

21. C. W. Belter, A Practical Guide for Informationists 33 (2018)

22. N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, Journal of Business Research **133**, 285 (2021)

23. T. Saheb, B. Amini, and F. Kiaei Alamdari, International Journal of Information Management Data Insights **1**, (2021)

24. D. Liu, J. Zhou, D. Pan, Y. Peng, and X. Peng, Measurement: Journal of the International Measurement Confederation **63**, 143 (2015)

25. X. Shu, G. Li, Y. Zhang, J. Shen, Z. Chen, and Y. Liu, Journal of Power Sources **471**, (2020)

26. C. Zhang, Y. He, L. Yuan, and S. Xiang, IEEE Access **5**, 12061 (2017)

27. W. A. Yang, M. Xiao, W. Zhou, Y. Guo, and W. Liao, Shock and Vibration **2016**, (2016)

28. X. Shu, G. Li, J. Shen, Z. Lei, Z. Chen, and Y. Liu, Energy **204**, (2020)

29. Y. SONG, D. LIU, Y. HOU, J. YU, and Y. PENG, Chinese Journal of Aeronautics **31**, 31 (2018)

30. C. S. Chin and Z. Gao, Energies **11**, (2018)

31. M. Attarian Shandiz and R. Gauvin, Computational Materials Science **117**, 270 (2016)

32. R. Yang, R. Xiong, W. Shen, and X. Lin, Engineering **7**, 395 (2021)

33. P. K. R. Maddikunta, G. Srivastava, T. R. Gadekallu, N. Deepa, and P. Boopathy, IET Intelligent Transport Systems **14**, 1388 (2020)

34. S. Sun, J. Zhang, J. Bi, Y. Wang, and M. H. Y. Moghaddam, Journal of Advanced Transportation **2019**, (2019)

35. W. Li, J. Zhu, Y. Xia, M. B. Gorji, and T. Wierzbicki, Joule **3**, 2703 (2019)

36. X. Tang, K. Yao, B. Liu, W. Hu, and F. Gao, Energies 2018, Vol. 11, Page 86 **11**, 86 (2018)