STM JOURNALS

International Journal of
Membranes

ISSN: 3049-4427
Volume 3, Issue 1, 2026
January–June
DOI (Journal): 10.37591/IJM

https://journals.stmjournals.com/ijm

Review IJM

Machine Learning for Finding Materials for Membranes

V. Basil Hans*

Abstract

Traditionally, finding and improving membrane materials has depended on trial-and-error experiments, which can take a long time, cost a lot of money, and only cover a small area. Recent improvements in machine learning (ML) have the potential to change the way membrane materials are designed by making it possible to make predictions about performance, selectivity, and stability based on data. ML algorithms can find hidden links between the structure, composition, and separation efficiency of membranes. This can help scientists design new materials for gas separation, water purification, and energy use in a logical way. Machine learning (ML) frameworks can quickly search across large chemical spaces to find the best polymer chemistries, nanoporous architectures, and mixed-matrix membrane topologies by combining computational modelling, high-throughput simulations, and experimental data. Additionally, the integration of machine learning with molecular dynamics and density functional theory improves comprehension of nanoscale transport mechanisms. Even if a lot of progress has been made, there are still problems with standardising data, making models understandable, and adding experimental feedback loops. To create strong, generalisable models that can lead to the next generation of high-performance, sustainable membrane technologies, these problems need to be solved.

Keywords: Machine learning, membrane materials, data-driven design, and molecular modelling Material informatics, predictive modelling, and separation performance

INTRODUCTION

Membrane separations, which are often used to separate gases and treat water, have become more popular as energy-efficient alternatives to traditional methods. Membrane materials for these uses need to be very permeable without losing selectivity, stable in the conditions they will be used in, and cheap to make. So, membrane development is necessary to make large-scale separation methods cost-effective.

Conventional experimental methodologies for identifying novel membrane materials depend on combinatorial chemistry or the prioritisation of the material production pathway; nevertheless, they are often sluggish and resource-intensive. Machine learning can assist cut down on the number of materials that need to be tested in experiments. Membrane-membrane performance indices that combine permeability and selectivity work well with diverse materials and let you use computers to screen prospective compositions without having to figure out all of their attributes [1]. Machine-learning models trained on a limited dataset of experimental membrane measurements can forecast performance indices for polymers and function under constrained design parameters, thereby indicating further possibilities for investigation. Machine-learning techniques help accelerate the predictive-design cycle for membranes with intricate formation chemistries that provide additional challenges to investigation [2].

This paper looks at the problems and chances that machine learning and other data-driven methods bring for speeding up the search for new membrane materials.

***Author for Correspondence**
V. Basil Hans
E-mail: vhans2011@gmail.com

Research Professor, Department of Commerce and Management, Srinivas University, Mangalore, Karnataka, India

Received Date: November 10, 2025
Accepted Date: December 13, 2025
Published Date: January 20, 2026

Citation: V. Basil Hans. Machine Learning for Finding Materials for Membranes. International Journal of Membranes. 2026; 3(1): 1–7p.

FUNDAMENTALS OF MEMBRANE SUBSTANCES

Membranes are barriers that let certain materials pass through them. In membrane sequence, how well the materials work depends on how permeable and selective they are. A material is called selective if it separates a certain blend of good and bad parts and works well at doing so. When thinking about how pore size, diffusion of component to pore aperture, fouling resistance, thermal and mechanical stability, and other factors affect membrane performance, we can look at common membrane materials. Taking this into account could greatly speed up the development of membrane materials. In the search for fossil energy and membrane materials, several different materials were thought to be gas separation membranes. Polymeric membranes have come a long way, from basic structures to coordination compounds and bio-mimetic structures.

Most of the time, people still employ the same membrane materials. It is essential to investigate membrane materials, as the fundamental data for the development and evaluation of new materials mostly involves layer separation and is commonly utilised for initial assessment throughout the early stages of membrane materials discovery; current data may be integrated. The three-stage sulfonated hydrocarbon copolymer membrane has great proton conductivity and stability, but it isn't very useful in fuel cells. We also need separate polystyrene sulfonated, PVA cross-linking, and phenolic resol without cross-linking. The advantages and disadvantages of the aforementioned membrane-training data guide the construction of the three-stage sulfonated hydrocarbon copolymer for the intended membrane design. The sample size is enormous, and the training reference is divided into groups. The compression data from Grignard functionalisation makes it possible to explore the design material with the desired functional group [3–5].

WHAT MACHINE LEARNING ADDS TO DISCOVERY

Machine learning can speed up the hunt for improved membrane materials and systems. Early materials screening finds links between composition, structure, and performance. However, it can be hard to choose candidates for experimental validation. Machine learning enables an additional methodology: constructing models that accurately forecast specific qualities from screening descriptors. These kinds of models quickly provide the best possibilities, which cuts down on expensive lab work. One study, for instance, used 62 simple structure characteristics to rank the effectiveness of 11,562 polymers as gas separation membranes. If the goal is to add other material families to the first set, the models can suggest particular new structures. This could lead to hybrid or fully new materials that nevertheless perform as expected.

Machine-learning methods include a wide range of jobs that help find new materials. The most common job in membrane research is prediction: figuring out one or more performance indicators based on a description of the material. The second most important job is screening, which means comparing the performance of different individuals and finding the ones with the most potential. The third duty is called optimisation, and it means making an original candidate better in order to make performance better. The fourth task is design: to come up with wholly new materials that are predicted to work well.

INFORMATION FOR FINDING MEMBRANES

To use machine learning to find membrane materials, it's vital to gather data on membranes, their properties, and how well they work. This part talks about the different kinds of data that are available, how to check their quality and whether they are good for machine learning, and what their forms and contents are.

Different public sources and publications put together information about different types of membrane materials. The Membrane Materials Database (MEMD) is a compendium of materials whose properties have been tested in the lab. It is one of the most dependable places to find useful information. This collection provides information about the structure, the arrangement of the membranes, the chemical formula, and the chemical parts. Some performance metrics are permeability, selectivity, stability, and cost. You can search by goal (such morphology, stability, separation, or synthesis), process type, or

feed component (like CO_2 or O_2). Each structure is carefully chosen from the literature to make sure that both the materials and the predictions are correct. Standard guidelines [2] are used to add notes to performance data. Other resources have similar information, however MEMD is mostly about membrane materials.

When getting a dataset ready for machine learning, you need to think about the features (also called descriptors) that characterise the input and the labels that tell you what the output should be. Membrane descriptors that describe the structure of a material are better at predicting how well it will work than descriptions of polymer characteristics. There must also be standards for labelling that spell out which membranes are in the collection, what feature is of interest (like permeability, durability, or processability), and what value was found (like average or maximum) [6]. The collection may only include data from membranes that are the same or from combinations that aren't fixed equipment, including membranes that are mixed into paints, separators, or catalysers.

Data for membrane materials often has a few problems that can make it harder to train machine learning models. It is common to come across incomplete datasets that only have a small number of the records you want. A lot of collections rely on bibliometric extraction and don't include works that haven't been published. A substantial percentage of stated performances either lack experimental validation or rely on second-hand methodologies. So, more materials should be collected because coverage is usually less than 100%. Bias may also influence the records, arising from publication patterns and readily available journals; for example, the majority of items in the MEMD concentrate on CO_2 and organic solvents. This kind of bias still lets you train on a wider range of materials, but it might not work well for diverse targets. Consequently, it is imperative to evaluate the bias and comprehensiveness of a current database and incorporate more data to facilitate the development of extendable prediction models. Noise can also mess with the data – misprints and other mistakes that happen when extracting literature might change the membrane's performance, characteristics, and even its identity. So, before using the set, it is important to fix these problems to make the materials more reliable.

WAYS TO MODEL

It can take a lot of trial and error to find materials with the best membrane properties. This can mean making and testing hundreds of candidates. Machine learning (ML) can help. It makes predictions about how well a material will work based on examples of other materials that have been evaluated. When you provide a trained model new candidates or chemical structures, it tells you which ones are the best to look at more closely [1].

There are three main types of machine learning methods for finding new membrane materials: regression, classification, and surrogate modelling. In regression, models guess at features that change over time, like permeability or selectivity. Classification systems put candidates in order based on labels like "medium" or "high" performance. Surrogate models estimate the results of complicated simulations, which makes it easy to quickly test structures [2]. It is common to check performance models by partitioning the data into training and test sets, although overfitting is still a problem.

MAKING ML WORKFLOWS THAT ARE EASY TO UNDERSTAND AND USE

By using simple workflows that make use of cheap, easy-to-generate data, machine learning can help uncover superior membrane materials. Membranes keep things apart in gas, liquid, or vapour form. The best membrane materials have a lot of permeability and selectivity, are stable, and don't cost too much. Finding membranes requires a lot of experimentation and synthesis of billions of possible materials. ML speeds things up by suggesting good choices based on the composition or structure of materials, preferably before they are made or tested. Some common ML jobs are predicting important properties, screening a lot of materials to find the best ones, optimising materials with the right mix of properties, and making design suggestions based on what is already known.

The normal approach of define–collect–prepare–train–test–interpret–propose works well with ML for membranes. The first step in defining the problem makes it clear what the query is and helps guide the next steps. Research on membranes usually falls into one of four main groups: gas separation, pervaporation, solvent-resistant nanofiltration (SRNF), or wastewater treatment. Each group has its own set of goal features, such as permeability, selectivity, and resistance to fouling. After that, you can put together the right datasets from public databases. Data preparation changes the raw data into a format that ML algorithms can use. This is generally done by making descriptors that show important aspects of the materials being considered. The training uses the desired attributes as goals. The rest of the established procedure focusses on building models, testing their performance, and finding better candidates [7].

MEMBRANE CASE STUDIES

Research aimed at mitigating global water scarcity has intensified interest in membrane filtration technologies. In this case, membranes are physical barriers that only let certain things (such water, ions, and organic molecules) through while keeping others out. This is similar to how a coffee filter keeps coffee grounds from getting into the drink. Membranes are used a lot in the food, pharmaceutical, semiconductor, metal, mining, and oil industries because they help save energy when cleaning water and recycling resources. Membrane-based systems can use up to 90% less energy than traditional separation methods. To find better membrane materials, you need to know a lot about how their structure and performance are related at both the molecular and macromolecular levels.

Traditional experimentation to investigate structure-performance correlations is expensive and time-consuming. Molecular modelling, molecular simulation, material characterisation, and property measurement are some of the basic inputs for peer-reviewed publications and patent applications. Material design procedures that use machine learning can find better candidate materials from property descriptors and speed up the process of getting candidates into real-world applications. Because of this, more and more people are interested in using machine learning to find promising membranes [3]. These kinds of workflows can help with laboratory synthesis by helping scientists choose which materials to test or by setting up the best synthesis processes. Batch materials, procedures, and equipment that are easy to use for automated experimentation can make things even more efficient.

A 2022 study outlined an approach to actively facilitate the search for prospective membranes, illustrated through a series of straightforward case studies focussing on various facets of membrane material discovery. The main purpose was to show the specific tactics used and the results gained in speeding up the process of designing membranes.

PROBLEMS AND LIMITS

Using Machine Learning to Find Membrane Materials

Membrane processes use barriers that let things through to keep things apart. Membranes are very important for water treatment, gas separation, food production, medicines, biotechnology, and energy storage. They help clean, separate, and classify things. Permeability, selectivity, stability, and cost are some of the most important qualities of a material. Machine learning helps scientists build on what they already know to identify better membranes faster. It also cuts down on the number of trials needed, makes predictions more accurate, and suggests new materials and changes that could lead to new discoveries.

Membrane materials can be made of polymers, ceramics, metals, or composites. Most of the time, polymeric membranes are employed. Polysulfone, polyethersulfone, polyimide, polyvinylidene fluoride, polyvinyl chloride, and polycarbonate are some of the most common types of membrane polymers. The performance of a membrane depends on the properties of the polymer, such as the size of the pores, how well it diffuses, how well it resists fouling, how stable it is at high temperatures, and how strong it is. Membranes are classified based on permeance, rejection, and other important factors using standard testing methods [2]. People share a lot of short reports that summarise a few important parameters. This is great data for machine learning.

THE BEST WAYS FOR TEAMS TO WORK TOGETHER

Teamwork is important for success with machine learning. Set clear roles and make a plan for how people from different fields can work together. Plan for personnel changes and maintain track of important information so that everyone on the team is up to date. Manage data throughout the process by setting up a common repository, making rules for changes and provenance, and writing up everything in detail.

To Make Your Machine Learning Work More Efficient and Useful, Follow These Recommended Practices

- Set obligations early on. Give people the job of modelling, software, data, and domain knowledge. Think about who will be on the team when you plan the project and write down the parameters for getting good input.
- Data for control. Make a shared place for everyone to work together and make sure things may be repeated. To keep track of changes and make sure they can be traced, use a version-control system. Give them training or templates for following the standards, keep track of changes, set up guidelines for validation, and make it clear what counts as complete and correct data.
- Write down important information. Keep track of the initial conditions, the status at hand-offs, the source of the training data, the models that were tested, the findings, and the actions that were taken after that. Use common words and phrases. Focus on information that is important for understanding.
- Share the code. Make scripts, notebooks, and other code easy to find, and include comments, examples, and tests when they might be useful. Allow others to get the same results without having to install additional software or figure out how to read proprietary formats. Use tools like GitHub, Jupyter, and Binder to make sharing easier and make sure everyone can see it.
- Use an iterative method. Come up with concepts that can be swiftly tested and changed based on what you learn. When there are several choices for things like picking data or choosing an algorithm, make and test more than one at the same time [2].

WHAT'S NEXT

Machine learning is making great strides, which bodes well for the future of discovering new membrane materials. These novel materials can be used in intriguing and already common ways, like purifying water, separating gases, or making fuel cells. Combining laboratory automation with machine learning and active learning methodologies makes it possible to explore design space more quickly and accurately. Future research can integrate emerging multimodal datasets in the design of membrane materials; for instance, the amalgamation of catalytic activity and solubility information can expedite the identification of multifunctional catalytic membranes for carbon capture and utilisation applications. Big businesses and new businesses are also working on membrane projects, which could lead to a good return on investment and a big influence on society [2].

Machine learning can help find new membrane materials faster and with fewer tests. It can also provide very accurate predictions without needing high-quality material databases. Membranes are important for a lot of things, such cleaning water, separating gases, making batteries, and making fuel cells. Conventional trials for membrane design are laborious and expensive, prompting researchers to explore potential possibilities through chemical graphs, molecular fingerprints, or alternative material descriptors. Active learning can help you pick the best candidates to get the most knowledge from just a few experiments.

There are a lot of mature uses for membrane technology, like sterile filtration, haemodialysis, water purification, and gas separation. There are also a lot of niche uses for successful membrane-based separation and processing of fluid mixtures [8].

The developing membrane technology combined with machine learning (ML) algorithms can greatly improve the design and production of new membrane materials. Several studies have documented ML-

based support in membrane formation; however, none have provided a comprehensive overview of current ML-assisted membrane production techniques from a material design standpoint [9].

Polymeric membranes have been the most popular type of RO desalination membrane so far. The research effort was mostly about finding the best polymeric membrane materials from the late 1950s to the 1980s. In the following decades, the efficacy of RO membranes has been enhanced through the regulation of membrane formation reactions and the incorporation of poly-condensation catalysts and additives [10].

CONCLUSION

Membrane materials are very important for separation. It used to take years and a lot of experiments to find new materials. Machine learning can help find novel materials by quickly estimating how well membranes will work and offering new materials that could work. There are four main types of membrane materials: polymeric, ceramic, metal-organic frameworks, and carbon-based materials. There are several types of materials in each class, and they all have variable levels of selectivity and permeability. You can get reliable information about many of these materials from the literature, which makes it easier to find new membranes based on data.

From chemical structures and process characteristics, machine-learning systems can guess how well a membrane will work. A convolutional neural network forecasting pure-water permeance indicates that the membrane structure significantly affects permeance, whereas process variables exert minimal influence. Another method relies solely on process characteristics to forecast membrane performance. For zeolite membranes, a careful neural network picks out the most important zeolite structure from a lot of data. A machine-learning algorithm can reliably estimate the concentration of the permeate based on the structure of the membrane, the shape of the feed-through, and the conditions under which it is running. These strategies can be made much more effective using active learning. Machine learning can speed up the discovery of membranes, which will assist fulfil the growing need for separation technologies in many areas .

REFERENCES

1. Krishnapriyan AS, Montoya J, Haranczyk M, Hummelshøj J, Morozov D. Machine learning utilizing persistent homology and chemical word embeddings enhances prediction accuracy and interpretability in metal–organic frameworks. Sci Rep. 2021 Apr 26;11(1):8888. doi: 10.1038/s41598-021-88027-8.
2. Glass S, Schmidt M, Merten P, Abdul Latif A, et al. Design of modified polymer membranes using machine learning. 2024. Available from: https://www.ncbi.nlm.nih.gov.
3. Maleki R, Shams SM, Mehdizadeh Chellehbari Y, Rezvantalab S, et al. Materials discovery of ion-selective membranes using artificial intelligence. 2022. Available from: https://www.ncbi.nlm.nih.gov.
4. Yu X. The use of machine learning in the life of materials. arXiv preprint arXiv:1707.04826. 2017. Available from: https://arxiv.org.
5. Oliveira ON, Oliveira MCF. Materials discovery with machine learning and knowledge discovery. 2022. Available from: https://www.ncbi.nlm.nih.gov.
6. Giro R, Hsu H, Kishimoto A, et al. AI-powered automated discovery of polymer membranes for carbon capture. npj Comput Mater. 2023;9:133. doi: 10.1038/s41524-023-01088-3.
7. Conrad F, Mälzer M, Schwarzenberger M, Wiemer H, et al. Benchmarking AutoML for regression tasks on small tabular data in materials design. 2022. Available from: https://www.ncbi.nlm.nih.gov.
8. Nunes SP, Culfaz-Emecen ZP, Ramon GZ, Visser T, Koops GH, Jin W, et al. Thinking the future of membranes: perspectives for advanced and new membrane materials and manufacturing processes. J Membr Sci. 2020 Mar 15;598:117761. doi: 10.1016/j.memsci.2019.117761.

9. Talukder MJ, Alshami AS, Tayyebi A, Ismail N, Yu X. Membrane science meets machine learning: future potential in assisting membrane material design and fabrication. Sep Sci Technol. 2023 May 11. doi: 10.1080/15422119.2023.2212295.
10. Lee KP, Arnot TC, Mattia D. A review of reverse osmosis membrane materials for desalination: development to date and future possibilities. J Membr Sci. 2011 Mar 15;370(1–2):1–22. doi: 10.1016/j.memsci.2010.12.036.

