JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SynthoChemAI: Deep Learning In Chemistry For Enhancing Synthetic Pathways Through AI-Based Reaction Prediction

Gurjant Singh[1]

[1] MSc. Botany & PGD in AI in Agriculture, Department of Botanical and Environmental Sciences, Guru Nanak Dev University, Amritsar, Punjab (INDIA)

ABSTRACT

"SynthoChemAI" is a compound term derived from "Synthesis," "Chemistry," and "Artificial Intelligence." It refers to the application of AI in the field of chemical synthesis, aiming to optimize processes, predict outcomes, and discover new chemical compounds. Synthetic chemistry is undergoing a revolution thanks to the integration of deep learning technologies and artificial intelligence (AI). This is especially true for reaction prediction. In comparison to conventional heuristic methods, this paper examines the notable developments in AI-driven models that improve prediction accuracy and efficiency of chemical reactions. These models speed up drug discovery and materials development by identifying the best reaction pathways by utilising large chemical databases like Reaxys and PubChem. We go over the state of deep learning applications today, including different model architectures and how well they work in practical situations. Nonetheless, there are still problems to be solved, such as interpretability of the model, data quality, and the requirement for integration with experimental chemistry. The paper also discusses upcoming trends that need to be taken into account as AI technologies proliferate, such as hybrid models that incorporate AI with traditional approaches and reinforcement learning. In the end, our work highlights the revolutionary potential of deep learning in synthetic chemistry, opening the door to more effective and inventive research methodologies that have the potential to have a big influence on materials science and global health.

Keywords: SynthoChemAI, Artificial Intelligence, Deep Learning, Reaction Prediction, Synthetic Chemistry, Drug Discovery, Chemical Databases, Reinforcement Learning.

I. INTRODUCTION

"SynthoChemAI" is a compound word that is made up of three essential parts. The word "**synthesis**," which describes the process of joining various components to form new compounds, especially in the realm of chemistry, is the source of the prefix "**Syntho**". "**Chem**" is an abbreviation for "**chemistry**," the scientific field that examines the characteristics, makeup, and actions of matter. Finally, "**AI**" is an acronym for "**Artificial Intelligence**," which refers to the process by which computers and technology mimic human intelligence.

*"**SynthoChemAI**" as a whole captures the nexus between chemical synthesis and artificial intelligence, emphasising its capacity to forecast results, optimise procedures, and aid in the identification of novel chemical compounds.*

As the basis for creating novel medications and cutting-edge materials, synthetic chemistry is essential to drug discovery and materials science. The efficient design and synthesis of new chemicals is a necessary skill for tackling intricate problems in technology and health. For example, it can take a lot of time and resources to synthesise physiologically active compounds since it involves precise control over chemical reactions and an understanding of reaction pathways (Campos, K.R. et al., 2019).

Artificial intelligence (AI) and deep learning technologies have become formidable instruments in synthetic chemistry in recent times, completely changing the way chemists approach compound design and reaction prediction. These systems use massive datasets of chemical reactions to find patterns and make highly accurate predictions about the results of new reactions (Niazi, S.K. et al., 2023; Valavanidis, Athanasios. 2023). Researchers can optimise reaction conditions and reduce the trial-and-error process usually involved with synthetic pathways by using AI-driven models that analyse a variety of chemical data, such as catalyst performance, reaction conditions, and structural properties (Shoichi Ishida, et al., 2022).

This review aims to investigate the ways in which AI-powered models are revolutionising reaction prediction in synthetic chemistry. This paper shows how machine learning algorithms can be used to anticipate reaction outcomes and optimise synthesis methods, thereby accelerating the speed of discovery in materials research and medicine development. We will also talk about the difficulties and restrictions associated with the use of AI in synthetic chemistry today, as well as potential future paths for this quickly developing field of study. [Fig. 1,2]

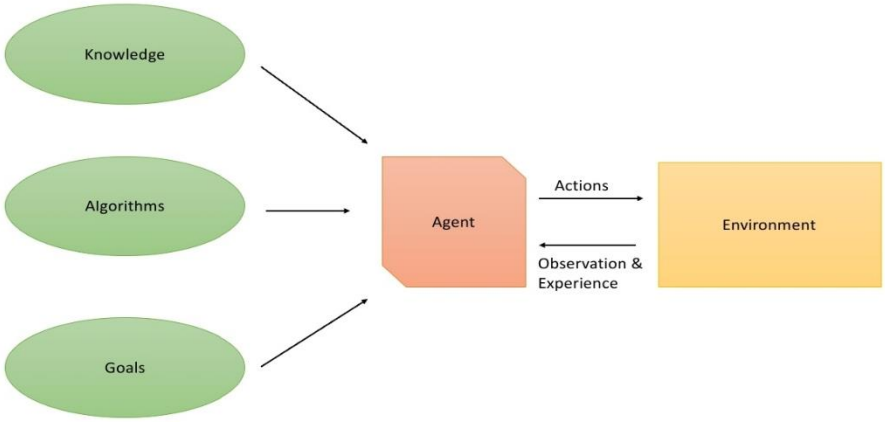


Fig. 1: Block Diagrams of an AI (Artificial Intelligence) System

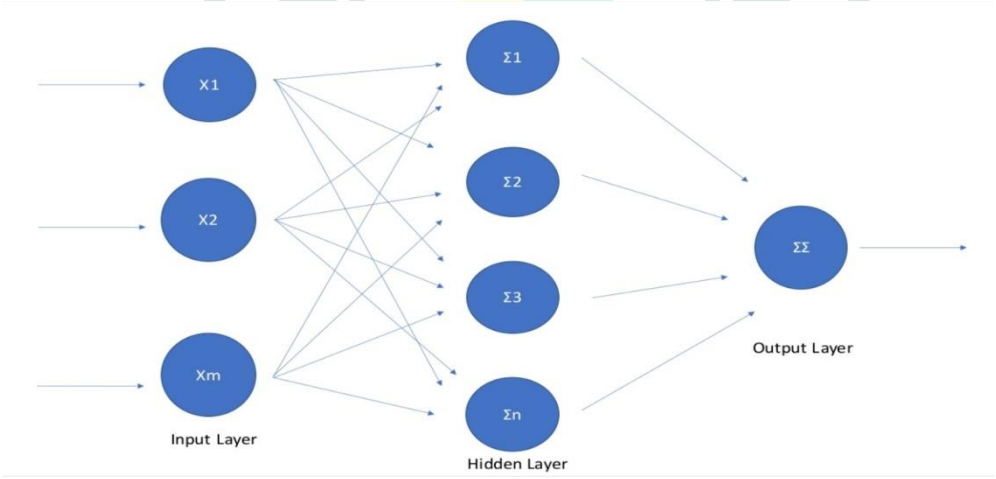


Fig. 2: Typical Architecture of Artificial Neural Network System

II. LITERATURE REVIEW

1. *Traditional Approaches to Reaction Prediction*

In the past, rule-based and heuristic approaches have been largely utilised in synthetic chemistry for reaction prediction. These conventional methods predict results using empirical criteria established from years of chemical research and well-understood reaction pathways. These approaches struggle with novel or complex reactions, which limits their scalability and adaptability even while they work well in well-characterized systems. Rule-based systems, such as expert systems and cheminformatics tools, are extensively utilised; yet, their application is limited due to the requirement for human skill and curation (Chengchun Liu, et al., 2024).

2. *Emergence of AI in Chemistry*

The development of deep learning models in particular, together with artificial intelligence, has created new opportunities for reaction prediction. Deep learning methods, such recurrent neural networks (RNNs) and convolutional neural networks (CNNs), have been used in chemistry recently, allowing models to learn patterns directly from reaction data without the need for pre-established rules. Significant improvements in predicted accuracy and the capacity to manage intricate reaction pathways have resulted from this change (Rizvi Syed Aal E Ali, et al., 2024; Han R. et al., 2023).

3. *Key Developments in Deep Learning for Reaction Prediction*

Recent advances have demonstrated that AI-driven models are capable of predicting reactions more accurately than traditional techniques. The invention of the Molecular Transformer by Schwaller et al. (2019), which makes use of a transformer-based neural network model, was a significant breakthrough. This model was able to predict reaction outcomes with previously unheard-of accuracy after being trained on a sizable reaction dataset. Other noteworthy models include DeepChem, which models chemical structures as graphs and uses graph-based deep learning approaches to predict reaction pathways (Kovács, D.P. et al., 2021).

4. *Reaction Databases and Data-Driven Approaches*

Both the volume and quality of data used during training determines how well AI models perform. Thanks to their vast collections of chemical data and reactions, databases like Reaxys, USPTO, and PubChem have grown to be indispensable resources. Journals, Stm et al., 2024) state that these datasets offer the variety of chemical environments required for building strong AI models. Problems with data quality still exist, though, since missing and inconsistent data can have a detrimental impact on model performance.

5. *Performance Metrics and Model Comparisons*

The effectiveness of traditional approaches and AI-based models has been examined in several research. According to Chen, LY. & Li, YP. (2024), deep learning models are more accurate and efficient at predicting chemical reactions, particularly when dealing with intricate, multi-step processes. AI models routinely beat rule-based systems in reaction prediction tasks, according to metrics like precision, recall, and F1-scores generated from confusion matrices (Lu, Jieyu & Zhang, Yingkai, 2022).

6. *Challenges and Limitations in Current Research*

AI models in reaction prediction encounter a number of difficulties despite their potential. Interpretability issues are brought up by deep learning models' "black box" nature. Chemists frequently favour models that, unlike existing AI techniques, allow predictions to be linked back to mechanical understanding. Moreover, models frequently rely on theoretical data and lack direct feedback from laboratory results, making the integration of experimental chemistry with AI predictions a substantial challenge (Tran, T.T.V. et al., 2003).

III. MATERIALS AND METHODOLOGY

Selection Criteria for Literature

We especially chose articles for this review that dealt with the use of AI-driven models—in particular, deep learning techniques—for reaction prediction in synthetic chemistry. Papers released between 2010 and 2024 were the main emphasis to guarantee inclusion of the most recent developments. The selection of studies was focused on their practical applications in materials science and pharmaceuticals, model correctness, and usefulness to reaction pathway design. Peer-reviewed research and noteworthy case studies that provided numerical data or illustrated novel techniques were given precedence.

Databases and Sources

Reaxys, PubChem, USPTO, ScienceDirect, Google Scholar, and other prestigious databases and scientific journals provided the literature for this review. These resources were picked because of their vast datasets on chemical reactions and artificial intelligence (AI)-powered synthetic chemistry research. In addition, pertinent research on reaction prediction models was looked up in prestigious journals like Nature, ACS Central Science, and Journal of Chemical Information and Modelling.

Inclusion and Exclusion Criteria

Included were just those papers that reported computational or experimental outcomes pertaining to AI-based reaction prediction. Research that only addressed theoretical models or that did not include quantitative evaluation measures (such as accuracy, precision, or recall) were not included. Studies that were exclusively based on conventional rule-based systems were also excluded in order to keep the focus on cutting-edge AI methodologies—that is, unless they were directly compared to AI approaches.

Method of Data Analysis

Based on how AI models were used in the chosen studies, a methodical analysis was conducted, paying particular attention to the accuracy, speed, and scalability of the application of the models. Metrics seen commonly in the literature, such as confusion matrices, precision, recall, and F1-score, were used to assess the performance of various deep learning models. To evaluate gains in predicting performance, comparisons between conventional and AI-based techniques were also emphasised.

Limitations

There are several restrictions, even though this review aimed to cover the most important developments in AI-driven reaction prediction. Bias may be introduced by differences in reaction types, dataset sizes, and reporting quality amongst research. Furthermore, the emphasis on contemporary research may obscure prior foundational studies that advanced reaction prediction approaches without the use of AI. [Fig. 3]

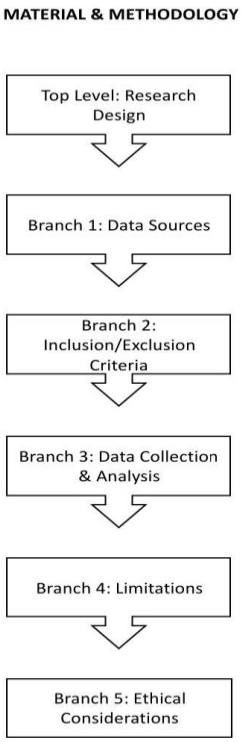


Fig. 3: Flow Chart Shows the Material & Methodology Employed for Review Article in Journal "Deep Learning In Chemistry: Enhancing Synthetic Pathways Through Al-Based Reaction Prediction"

IV. BACKGROUND ON REACTION PREDICTION

A. Definition of Reaction Prediction and Its Importance in Synthetic Pathways

The process of estimating a chemical reaction's products based on Its reactants, surroundings, and other pertinent variables is known as reaction prediction. It is an essential component of synthetic chemistry since it enables scientists to predict reaction outcomes, which simplifies the design of synthetic pathways and boosts compound synthesis efficiency (Tu Z, Stuyver T, Coley CW. 2022). The time and resources needed for experimental validation can be greatly decreased with accurate reaction predictions, accelerating the development of novel medications and materials. For instance, in the context of drug creation, the ability to forecast probable metabolic routes and adverse effects can help chemists create safer and more potent medications (Vamathevan, J. et al., 2019).

B. Traditional Methods of Reaction Prediction

Reaction prediction has historically depended on conventional techniques like rule-based systems and heuristics. Heuristics are methods of predicting reaction outcomes by applying empirical principles that are drawn from previous chemical knowledge and experimental data. For example, chemists frequently estimate potential products using well-established reactivity patterns, such as electrophilic and nucleophilic behaviour (Venkatasubramanian, Venkat & Mann, Vipul. 2021). On the other hand, rule-based systems make use of pre-established templates and rules to forecast reaction products in light of changes in functional groups and structural similarities. These techniques may not be as flexible as needed for more complicated reactions, but they can nevertheless effectively direct chemists in simple circumstances.

C. Limitations of Traditional Methods in Complex Reactions

Although conventional techniques for reaction prediction have proven useful, they have considerable drawbacks, especially when dealing with intricate reactions that involve several phases, a variety of functional groups, and fluctuating reaction circumstances. Their reliance on predetermined rules is one of their main drawbacks, as it might cause oversimplification and ignore particular reaction mechanisms (Ross, John, 2008). Furthermore, the investigation of new synthetic routes may be hampered by the inability

of conventional methods to accurately anticipate the results of reactions with unclear or poorly understood processes. Consequently, conventional techniques might not fully capture the complexities of contemporary synthetic issues, calling for the creation of more sophisticated predictive models that make use of AI and machine learning technology. [Table. 1]

Table. 1: Table covers the definition, traditional methods, and their limitations, emphasizing the shift towards AI models.

SELECTION	SUMMARY	REFERENCES
Definition of Reaction Prediction	To maximise compound synthesis and shorten experimental times, predict chemical reaction products to help with synthetic routes and medication development.	(Tu Z, Stuyver T, Coley CW. 2022; (Vamathevan, J. et al., 2019)
Traditional Methods	Based on templates and empirical guidelines (heuristics). utilised functional group shifts and reactivity patterns, but restricted in complex reactions.	(Venkatasubramanian, Venkat & Mann, Vipul. 2021)
Limitations	The necessity for AI-driven models for contemporary synthetic difficulties stems from the fact that conventional methods frequently oversimplify and fail in complex reactions or unclear mechanisms.	(Ross, John, 2008)

V. DEEP LEARNING TECHNIQUES IN REACTION PREDICTION

A. Overview of Deep Learning and Its Applications in Chemistry

Artificial neural networks are used in deep learning, a kind of machine learning, to model intricate correlations seen in data. It has been widely used in many domains, including chemistry, where it is used to analyse large datasets and forecast the behaviours of molecules and the results of reactions. Deep learning algorithms are able to identify patterns that conventional methods might miss because of their capacity to learn hierarchical representations from unprocessed data (Tu Z, Stuyver T, Coley CW. 2022). Deep learning has several uses in chemistry, from developing new compounds and optimising synthetic methods to forecasting molecular characteristics and reaction yields. These developments are especially helpful for catalysis, materials science, and drug discovery since they can lower costs and speed up the development process by forecasting reaction outcomes (Cova TFGG., Pais, AACC. 2019). [Fig. 4]

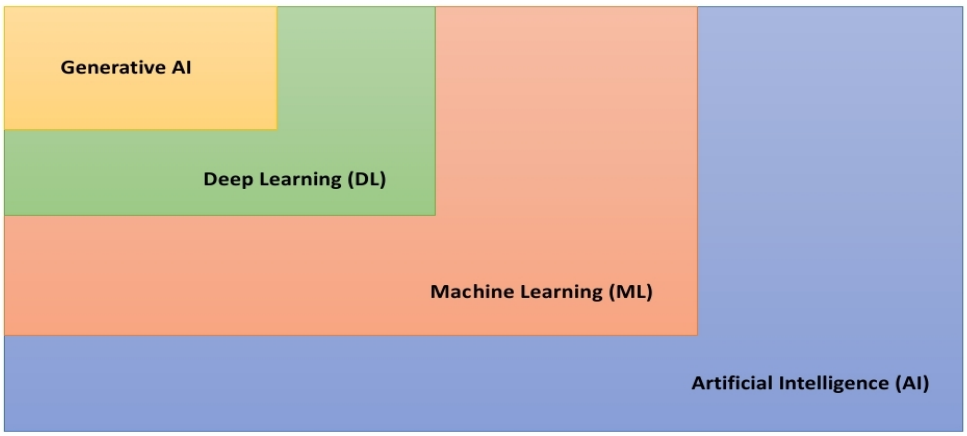


Fig. 4: Systematic Position of Deep Learning (DL) in the Evolution of AI

B. Types of Deep Learning Models Used for Reaction Prediction

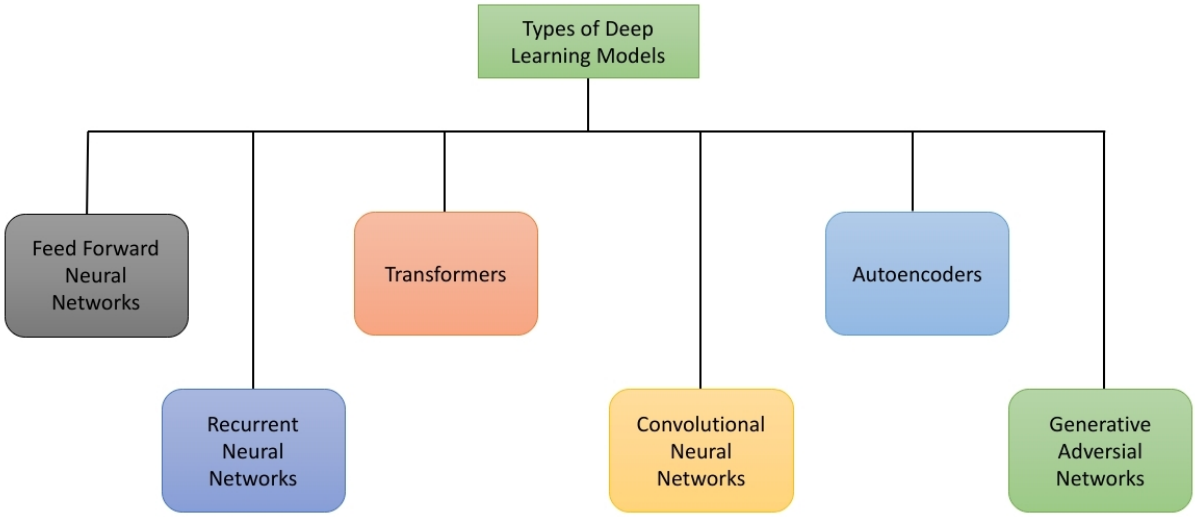
Various deep learning models have been effectively utilised for reaction prediction; the specific benefits of each model vary based on the type of data and the intricacy of the reactions being predicted. These models include:

1. Convolutional Neural Nets (CNNs): CNNs are mainly utilised for picture data, however they have been modified to represent chemicals by considering molecular structures as grid-like data or images. With this method, CNNs may efficiently predict reaction outcomes and automatically extract characteristics linked to molecular patterns (Hirohara, M. et al., 2018).

2. Recurrent Neural Networks (RNNs): RNNs are useful for modelling reaction mechanisms and forecasting multi-step reactions since they are especially well-suited for sequential data. Reactive neural networks (RNNs) are able to learn the dependencies between reactants and products over time by retaining knowledge from earlier steps in a sequence (Francesca Grisoni, et al., 2018).

3. 1. Graph Neural Networks (GNNs): Because GNNs can represent molecular structures as graphs with bonds acting as edges and atoms as nodes, they have become a potent tool for reaction prediction. Because of this structure, GNNs are very good at predicting the results of reactions by capturing the local and global context of molecular interactions (Besharatifard, M., Vafaee, F. 2024).

Fig. 5: Different Types of Deep Learning Models in AI



C. Discussion of Algorithms and Frameworks

The following frameworks and methods make it easier to use deep learning for reaction prediction:

1. *Chemoinformatics:* Chemoinformatics uses computational methods to analyse chemical data by combining computer science and chemistry. By using massive databases of chemical reactions for training, like the Reaxys or PubChem databases, deep learning models can be combined with chemoinformatics to improve the accuracy of reaction predictions (Wegner, Joerg 2012).

2. *Graph-Based Models:* As was already indicated, there is a growing trend in reaction prediction for graph-based models. More complex predictions are possible thanks to these models' ability to accurately depict chemical structures and their interactions. Implementing GNNs and other graph-based learning strategies is made easier with the help of frameworks like PyTorch Geometric and DeepChem (Jiang, D. et al., 2020).

3. *Transfer Learning:* Another intriguing strategy is transfer learning, which enables models developed for one job to be optimised for a related task. This helps researchers to employ current models for new applications, especially in chemistry where datasets may be restricted to particular types of reactions (Zhang, C. et al., 2024).

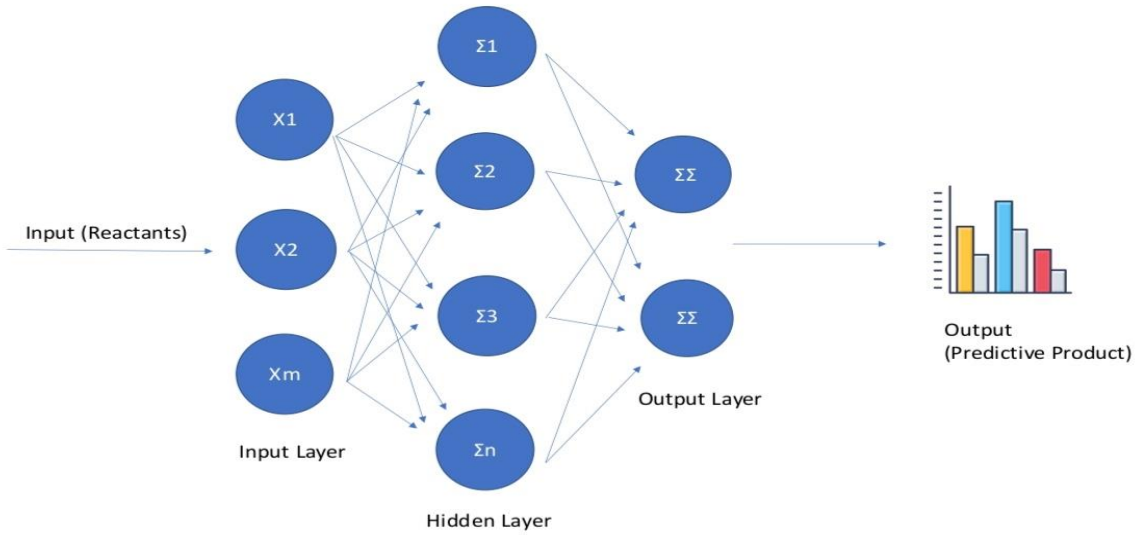


Fig. 6: Flow Diagram shows Simplified Version of the Mode of Action of ANNs for Reaction Prediction

VI. AI-BASED REACTION PREDICTION MODELS

A. Case Studies of Successful AI Models in Predicting Chemical Reactions

Reaction prediction has undergone a revolution thanks to artificial intelligence, which has produced a number of effective models with impressive accuracy and efficiency. These models use large chemical databases to train algorithms that require little assistance from chemists in order to predict reaction outcomes.

1. Overview of Databases Utilized for Training Models

Selected chemical databases play a major role in providing the data required for training and validation of AI models. The two most well-known databases are:

Reaxys: Reaxys Comprehensive information on chemical processes, including yields, product structures, and experimental conditions, is available in this database. It is a good resource for training AI models to predict response outcomes because it contains more than 30 million reactions (Lawson, Alexander et al., 2014).

PubChem: Maintained by the National Centre for Biotechnology Information (NCBI), PubChem is a popular database. It includes comprehensive details about chemical compounds, including their biological characteristics and activities. A wide range of AI applications are supported by the substantial amount of data accessible in PubChem, which allows models to learn from a variety of chemical interactions (Kim, S. et al., 2016).

2. Notable Models

Significant progress has been made by a number of AI models in the area of reaction prediction:

Reaction Transformer: To anticipate the end products of chemical processes, this model makes use of the transformer architecture, which was first created for natural language processing. The Reaction Transformer can efficiently learn from big datasets and produce precise predictions, even for intricate multi-step reactions, by encoding reactants as sequences (Zheng, Shuangjia, et al., 2019).

DeepChem: An open-source library called DeepChem combines chemoinformatics and deep learning. It has a variety of pre-trained models for different kinds of tasks, such as drug discovery, molecular property prediction, and reaction prediction. DeepChem makes AI modelling accessible to chemists and researchers by enabling rapid model construction and deployment (Ginsburg, Boris et al., 2020).

B. Comparison of AI Models with Traditional Methods in Terms of Accuracy and Efficiency

Artificial intelligence (AI) models have proven to perform better than conventional approaches for reaction prediction. AI models may learn directly from data, which enables them to capture complicated relationships and predict reactions with more accuracy than heuristics and rule-based systems, which frequently rely on empirical knowledge and predetermined rules (Struble TJ. Et al., 2020). Research has demonstrated, for example, that AI models can attain prediction accuracy levels of more than 90%, but conventional techniques can find it difficult to attain comparable levels, especially for intricate or uncommon reactions (Li, B. et al., 2023). AI models considerably cut down on the amount of time needed for reaction prediction in terms of efficiency. Conventional techniques can include a great deal of trial-and-error experimentation, whereas AI models may produce predictions in a matter of seconds, allowing for quicker decision-making and the investigation of novel synthetic pathways (Youngjun, Xu., et al., 2021).

C. The Role of Data Quality and Quantity in Model Performance

The calibre and volume of training data have a major impact on how well AI-based reaction prediction models work. Models are able to learn strong patterns and produce precise predictions across a range of reaction types when they are trained on high-quality, diverse datasets. On the other hand, inadequate or poor quality data could cause overfitting and poor generalisation to novel reactions (Chen, LY. et al., 2024). Furthermore, the training data's representation of molecular structures has a major influence on the model's performance. The model's capacity to comprehend and forecast chemical interactions can be improved by using standardised representations, such as graph-based representations or SMILES (Simplified Molecular Input Line Entry System) (Wu JN. et al., 2024). Therefore, progress in the area of AI-driven reaction prediction depends on efforts to curate and standardise chemical libraries. [Fig. 7]

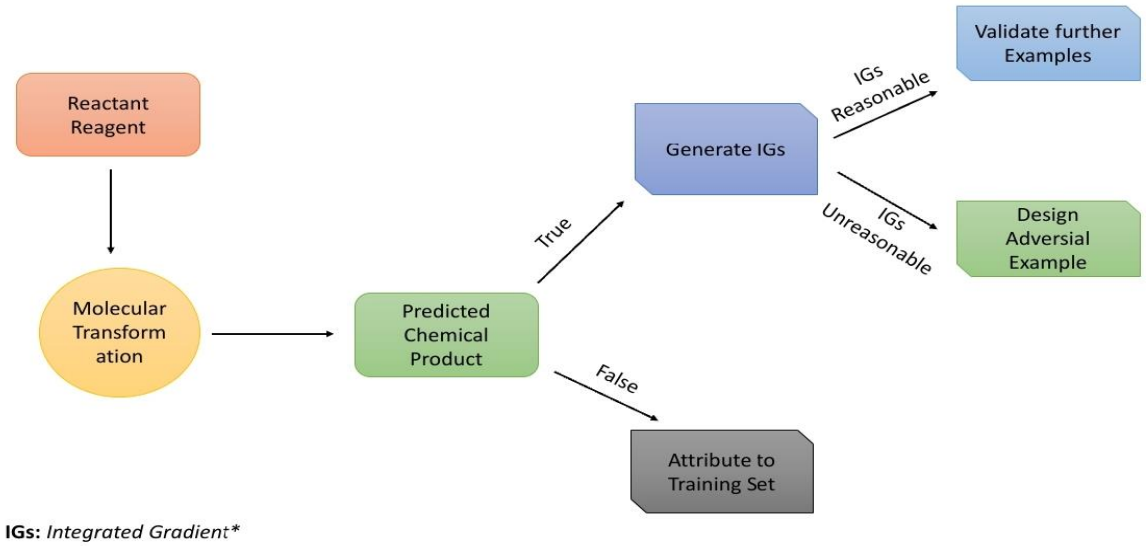


Fig. 7: Flow Diagram Depicting Systematic Process of AI-Assisted Chemical Reaction Prediction for Various Deep Learning Models.

D. Python Script: Reaction Prediction using Deep Learning

The Python script that follows can be used as a guide for a number of projects involving the use of AI and deep learning methods to analyse chemical interactions. Using a hypothetical dataset, this script shows how to build a basic machine learning model for reaction prediction. It makes use of well-known libraries like as TensorFlow (Abadi, M. et al., 2016), scikit-learn (Pedregosa, F. et al., 2011), and pandas (McKinney, W. 2010).

```
Import pandas as pd

Import numpy as np

From sklearn.model_selection import train_test_split

From sklearn.preprocessing import LabelEncoder
```

```
From tensorflow.keras.models import Sequential
From tensorflow.keras.layers import Dense, Dropout
From tensorflow.keras.utils import to_categorical
# Load the dataset
# Replace 'your_dataset.csv' with the actual dataset file path
Data = pd.read_csv('your_dataset.csv')
# Display the first few rows of the dataset
Print(data.head())
# Assuming the dataset has columns: 'reactants', 'products', and 'reaction_type'
# Preprocessing the data
# Encode categorical variables (reactants and products) to numerical values
Label_encoder = LabelEncoder()
Data['reactants'] = label_encoder.fit_transform(data['reactants'])
Data['products'] = label_encoder.fit_transform(data['products'])
# Features and target variable
X = data[['reactants']]
Y = data['reaction_type']
# Convert target variable to categorical format
Y = LabelEncoder().fit_transform(y)
Y = to_categorical(y)
# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
# Define the neural network model
Model = Sequential()
Model.add(Dense(64, activation='relu', input_shape=(X_train.shape[1],)))
Model.add(Dropout(0.2))
Model.add(Dense(32, activation='relu'))
Model.add(Dropout(0.2))
Model.add(Dense(y.shape[1], activation='softmax'))
# Compile the model
Model.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
# Train the model
Model.fit(X_train, y_train, epochs=100, batch_size=32, validation_split=0.1)
# Evaluate the model on the test set
Loss, accuracy = model.evaluate(X_test, y_test)
Print(f'Test Loss: {loss}, Test Accuracy: {accuracy}')
# Making predictions
# Example: Predict the reaction type for a new reactant
New_reactant = pd.DataFrame({'reactants': ['New Reactant Example']})
```

```
New_reactant['reactants'] = label_encoder.transform(new_reactant['reactants'])

Prediction = model.predict(new_reactant)

# Decode the predicted reaction type

Predicted_type = np.argmax(prediction, axis=1)

Predicted_type = label_encoder.inverse_transform(predicted_type)

Print(f'Predicted Reaction Type: {predicted_type}')
```

The Python script that is supplied can be used as a basic guide to apply deep learning methods to reaction prediction in synthetic chemistry. The pandas package is used to load a fictitious dataset at first; users are advised to substitute "your_dataset.csv" with the path to the real dataset. The script then uses Label Encoder to preprocess the data by turning category variables—like reactants and products—into numerical values.

It then defines the goal variable (y) and the characteristics (X), converting y into a categorical format appropriate for classification tasks. To enable model evaluation on unseen data, the dataset is split using train_test_split from sklearn into training and testing sets. Then, using TensorFlow, a neural network is built that consists of a hidden layer, an input layer with 64 neurones, and an output layer that uses the softmax activation function to predict the kinds of reactions is the result of a hidden layer having 32 neurones, dropout layers for regularisation, and an output layer (Abadi, M. et al. 2016)

To assess the model's performance, a portion of the training data is used for validation once the model has been assembled and trained on it. Following training, the test dataset is used to assess the model's performance and report accuracy and loss measures. The script concludes by demonstrating how to create predictions using fresh data. Note that this script is meant to be used as a basic guide, assuming that there is a training dataset that meets the requirements. It also highlights the need to install the necessary libraries, which are pandas (McKinney, W. 2010), NumPy (Oliphant, T. E. 2006), scikit-learn (Pedregosa, F. et al., 2011), and tensorflow (Abadi, M. et al., 2016). Users should modify the script in accordance with the particular requirements of their research as well as the structure of their dataset. [Fig. 8]

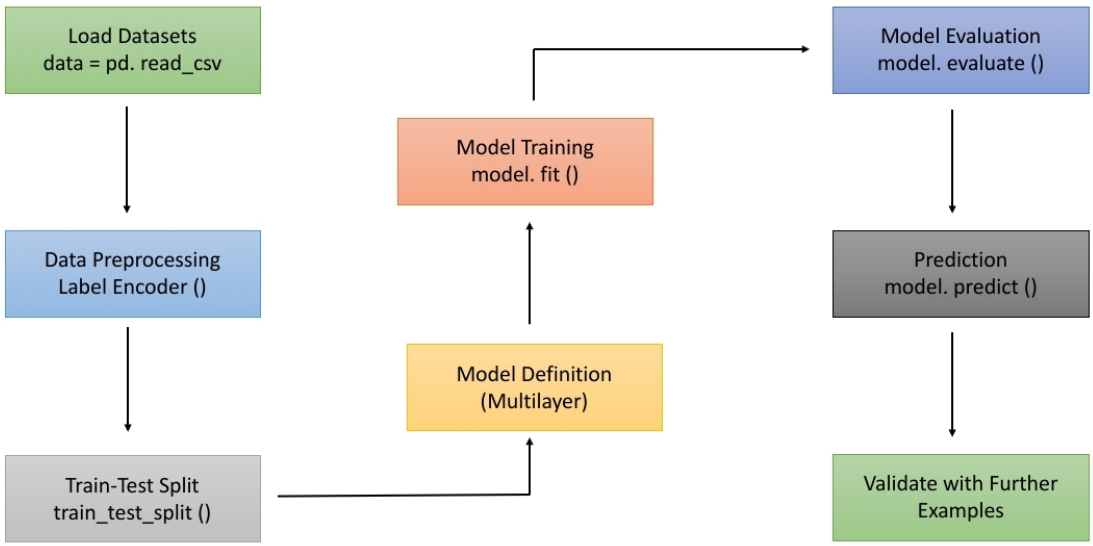


Fig. 8: Simplified Explanation of Python Programs for the Chemical Reaction Prediction using Deep Learning.

E. Performance Evaluation Tool

An essential tool for assessing the effectiveness of AI-based reaction prediction models is a confusion matrix. It offers a thorough analysis of the classification outcomes of the model, differentiating between various reaction types' true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Researchers may compute critical metrics like accuracy, precision, recall, and F1-score—all of which are essential for evaluating how well AI models predict chemical reactions—by arranging this data using the confusion matrix. [Fig. 9]

For example, a confusion matrix was used to assess how well a neural network model trained on big reaction datasets like Reaxys and USPTO performed in a study by Rafael, Gozalbes, et al. (2009). They were able to assess the model's prediction accuracy for different reaction classes using the confusion matrix, which also showed where the model worked best (high TP) and where it had difficulties (greater FN or FP). In a similar vein, Borrelli, William & Schrier, Joshua. 2021) evaluated the accuracy and recall of their transformer-based model using a confusion matrix, proving its potency in forecasting artificial chemical reactions.

In order to increase prediction accuracy and optimise algorithms, researchers can discover places where the model may over- or under-predict specific reaction types. This is made possible by the confusion matrix. Furthermore, it is necessary to guarantee that the AI model generalises well across a variety of datasets, improving its practical application in materials research and drug discovery (Pettit, R.W. et al., 2021).

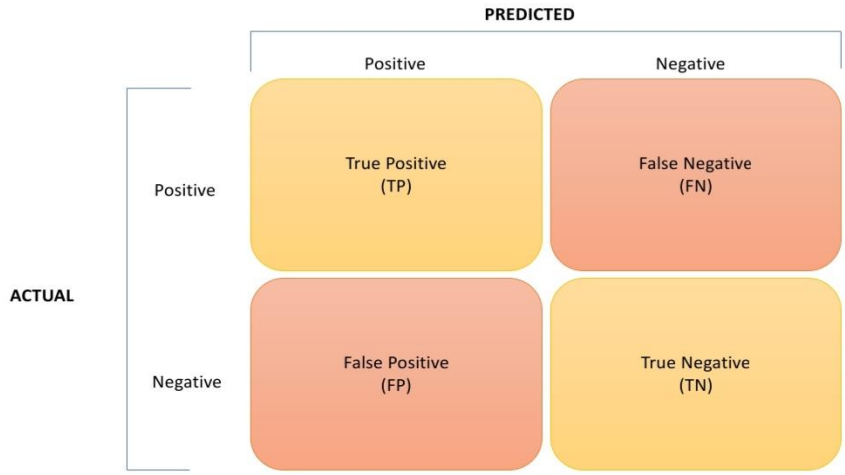


Fig. 9: Confusion Matrix for the Performance Evaluation of AI based Reaction Prediction Models.

1. **Accuracy:** Evaluates how accurate the model's predictions are overall.

$$Accuracy = TP + TN/TP + TN + FP + FN$$

2. **Precision**: Percentage of accurate positive predictions made for a specific response class.

$$Precision = TP_A/TP_A + FP_A$$

3. **Recall:** Evaluates the model's accuracy in identifying real positive cases.

$$Recall_A = TP_A/TP_A + FN_A$$

4. **F1-Score:** The harmonic mean of precision and recall.

$$F1_A = 2\times Precision_A \times Recall_A/Precision_A + Recall_A$$

TABLE 2: Comparison of model performance over four separate groups of our in-house dataset. The values of R^2 (R-Squared), MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) refers to the mean standard deviation across the folds (Li, B., Su, S., Zhu, C. et al. 2023).

GROUP	SIZE	METHODS	R^2	MAE	RMSE
G1	317	**Yield-BERT**	**0.712 ± 0.070**	**0.07 ± 0**	**0.10 ± 0.01**
G1	317	DeepReac+	0.544 ± 0.128	0.09 ± 0.01	0.13 ± 0.02
G2	419	**GraphRXN-sum**	**0.590 ± 0.034**	**0.06 ± 0**	**0.07 ± 0**
G2	419	Yield-BERT	0.512 ± 0.046	0.06 ± 0	0.08 ± 0.01
G3	401	**GraphRXN-concat**	**0.800 ± 0.030**	**0.06 ± 0**	**0.08 ± 0**
G3	401	DeepReac+	0.744 ± 0.032	0.07 ± 0.01	0.09 ± 0.01
G4	421	GraphRXH-sum	0.405 ± 0.091	0.09 ± 0.01	0.12 ± 0.01
G4	421	**Yield-BERT**	**0.490 ± 0.055**	**0.08 ± 0.01**	**0.11 ± 0.01**

*Bold represents the best fit model performance in each group.

VII. APPLICATIONS AND IMPACT ON SYNTHETIC CHEMISTRY

A. Enhancements in Reaction Pathway Design and Optimization

Design and reaction route optimisation have been improved as a result of the synthetic chemistry field's incorporation of AI and deep learning approaches. Chemists can systematically investigate a wide range of potential reaction conditions, catalysts, and substrates to find the best paths for target compounds by utilising sophisticated predictive models. For example, chemists can avoid difficult steps in conventional synthesis procedures by using AI models to propose alternate synthetic paths that may not have been previously investigated (Ali, Rizvi et al., 2024). Researchers may now access a larger chemical space because of this optimisation technique, which not only improves chemical synthesis efficiency but also opens up new pathways and reactions. [Fig. 10]

B. Real-World Applications in Pharmaceuticals and Materials Development

Applications of AI-driven reaction prediction models in materials science and pharmaceuticals are significant. These models aid in the quick discovery of possible drug candidates and the streamlining of artificial production processes in the pharmaceutical industry. AI algorithms have been utilised, for instance, to create novel antibiotics and anti-cancer drugs, greatly speeding up the process of discovery (Ali, K.A. et al., 2024).

Artificial intelligence (AI) models are used in materials research to anticipate and optimise interactions to produce sophisticated materials with specific features. Reaction prediction can be used in the synthesis of polymers, nanomaterials, and composites to determine the ideal conditions for obtaining specific properties like strength, conductivity, or biocompatibility (Champa-Bujaico E. et al., 2022). Furthermore, by optimising reactions that use greener solvents and reduce waste, AI-driven techniques can help in the production of sustainable materials. [Fig. 10]

C. Potential to Reduce Time and Costs Associated with Experimental Trials

The ability of AI-based reaction prediction models to cut down on the expenses and duration of experimental trials is one of its most important applications. Extensive trial-and-error experimentation is a common practice in traditional synthetic chemistry, which can be resource- and time-intensive. Chemists can concentrate their efforts on the most promising synthetic routes by prioritising experiments based on model predictions, which can be achieved by employing AI to anticipate reaction outcomes (Strieth-Kalthoff, et al., 2022).

Research has indicated that the use of AI-driven approaches can result in a 30% reduction in the total amount of time needed for the drug discovery process (Paul, D. et al., 2021). Furthermore, fewer experimental trials and less material waste translate into reduced overall expenses, increasing the economic viability of research and development. In the end, artificial intelligence (AI) in synthetic chemistry not only increases output but also supports sustainability objectives by reducing the environmental impact of chemical research (Mochen, Liao, et al., 2021). [Fig. 10]

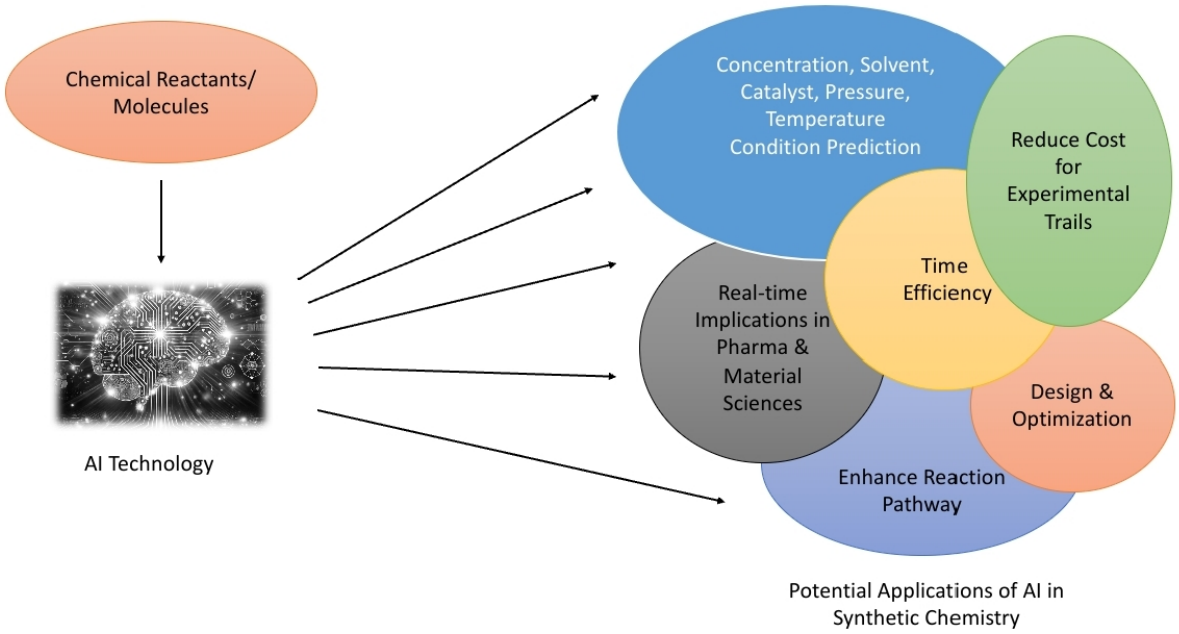


Fig. 10: Another Potential Applications like Prediction of Ideal Condition for Chemical Reaction.

VIII. CHALLENGES AND LIMITATIONS

A. Issues of Model Interpretability and the "Black Box" Problem

The problem of interpretability is one of the biggest obstacles to using AI models for reaction prediction. Many deep learning models, especially those that make use of intricate structures like neural networks, frequently function as "black boxes," meaning that users are not always able to easily understand the decision-making processes that go into creating their predictions (Carvalho, Diogo V. et al., 2019). For chemists who must validate and accept the predictions given by AI systems, this lack of transparency presents a major obstacle (Shah, Varun & Konda, Sreedhar. 2021). Understanding the reasoning behind predictions is essential for future investigation and experimental validation, and it might restrict a model's adoption in the scientific community if it is difficult to grasp why a model predicts a specific reaction pathway or result. [Fig. 11]

Scholars are presently investigating techniques to enhance the comprehensibility of these models. Chemistry researchers can better understand how input data affects predictions by using techniques like feature importance analysis, explainable AI frameworks, and attention processes, which attempt to illuminate the decision-making process of AI models. However, striking a balance between interpretability and model complexity continues to be a significant issue in the field (Singh, Shashank et al., 2024).

B. Challenges in Data Availability and Diversity

The availability and variety of training data is a key drawback for AI-based reaction prediction models. Although there is a lot of information available in big chemical databases like Reaxys and PubChem, there are differences in the quality and scope of this information. The generalisability of the models trained on these datasets is limited due to the fact that many of them are biassed towards particular reaction types or classes of substances (Tetko, I.V. et al., 2016).

Moreover, uncommon or unique reactions could not be adequately represented in the datasets that are currently accessible, which makes it difficult for AI models to predict results accurately in these situations. Because of this, models often perform remarkably well on well-studied reactions but struggle with more complicated or uncommon reaction types. The scientific community must

continue to collaborate in order to share data and insights, and to curate and grow chemical databases in order to address these data restrictions (Tran, T.T.V. et al., 2023). [Fig. 11]

C. Need for Integration with Experimental Chemistry

AI-driven reaction prediction models need to be combined with experimental chemistry procedures in order to function as intended. Although these models can indicate possible courses of action and results, empirical validation is still necessary to verify their accuracy. For models to be improved and their predictive power increased, an iterative feedback loop between computational predictions and experimental outcomes is essential (Shoichi Ishida, et al., 2022).

Additionally, chemists need to be trained in both computational and experimental methods in order to be able to evaluate and implement AI predictions in a laboratory context. In addition to improving the accuracy of AI-driven predictions, this combination of computational and experimental methods will promote creativity in synthetic chemistry by allowing chemists to investigate novel reactions and refine current procedures in light of well-informed forecasts (Han, R. et al., 2023). [Fig. 11]

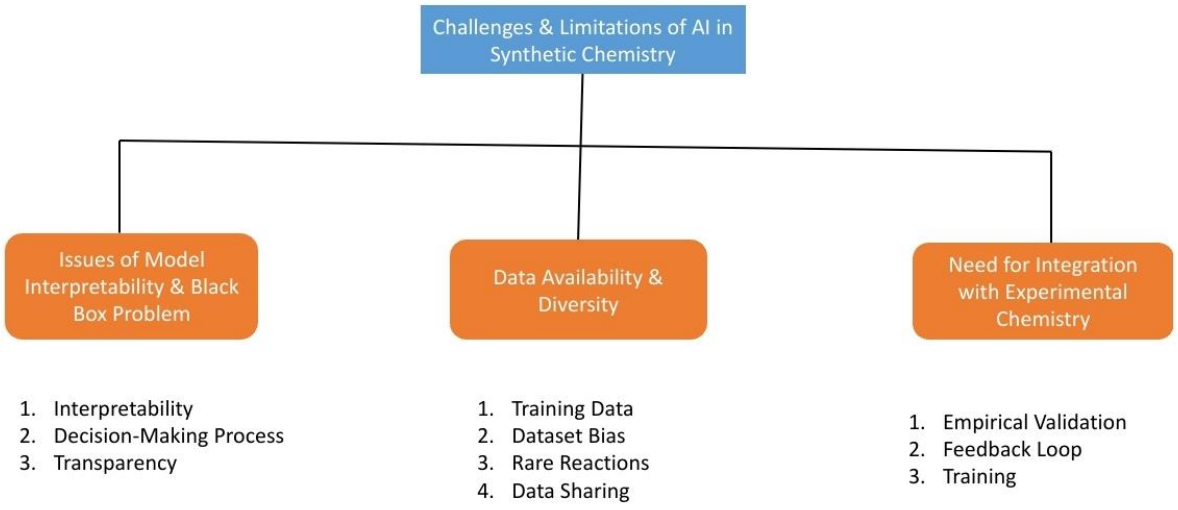


Fig. 11: Challenges & Limitations of AI in Synthetic Chemistry.

IX. FUTURE DIRECTIONS

A. Trends in AI and Machine Learning in Chemistry

AI and machine learning technologies are being rapidly incorporated into the area of chemistry, and a number of new trends are anticipated to have a significant impact on the discipline's future. The growing use of reinforcement learning (RL) in synthetic chemistry is one noteworthy trend. The optimisation of complex chemical reactions could be achieved dynamically through reinforcement learning (RL), a sort of machine learning where agents learn to make decisions by maximising cumulative rewards. Reactive learning (RL) is a technique that may be used to automate the process of designing experiments by continuously learning from trial-and-error interactions with chemical systems to identify optimal reaction pathways and circumstances (Zhenpeng Zhou et al., 2017).

Furthermore, developments in transfer learning are getting more and more attention; these developments enable models that are trained on one dataset to be converted to another with little further training. This can greatly improve the accuracy of reaction prediction, particularly when there is a lack of available data (Zhang, C. et al., 2024). The combination of AI and more advanced experimental methods will probably usher in a new era of chemical discovery that allows for the quick exploration of reaction space as computing power keeps rising.

B. The Potential of Hybrid Models Combining AI with Traditional Approaches

An intriguing prospect for improving reaction prediction and synthetic chemistry overall is the incorporation of AI with conventional techniques. The advantages of both methods can be combined in hybrid models that integrate heuristics and well-established

chemical principles with machine learning techniques. For instance, by directing the model based on established chemical behaviour, combining AI with mechanistic understanding can result in more accurate predictions (Bengani, Vedika. 2024).

Hybrid models can also aid in bridging the gap between computational and experimental chemistry. Researchers can learn more about reaction processes, improve conditions, and create predictive models that are comprehensible and supported by science by applying AI to evaluate data gathered from conventional experimental techniques (Visan AI. & Negut I. 2024). This combination may make it easier to comprehend complex processes more thoroughly and encourage the development of novel synthetic techniques.

C. Ethical Considerations and the Future Landscape of AI in Synthetic Chemistry

Ethics must be taken into account as artificial intelligence (AI) technologies continue to dominate the field of synthetic chemistry in order to ensure its appropriate and equitable application. Data privacy, the possibility of bias in AI models, and the effects of automating chemical discovery procedures are important concerns. Building equitable and successful models requires ensuring that datasets are representative and devoid of bias (Hermann, E. et al., 2021).

Moreover, the growing prevalence of AI-driven techniques raises questions about the ramifications for the chemical sciences workforce. Artificial intelligence (AI) has the ability to save expenses and increase production, but it also raises concerns about the future of chemistry and the possible loss of traditional jobs. To optimise the advantages of these technologies while maintaining job security and ethical responsibility in research and development, human chemists and AI systems must collaborate (Elendu, C. et al., 2023).

In summary, artificial intelligence (AI) in synthetic chemistry has a promising future that will be marked by advancements in hybrid models, reinforcement learning, and a growing emphasis on moral issues. Chemical research and development can become more inventive, efficient, and socially responsible by overcoming these obstacles and embracing new technology.

X. CONCLUSION

This paper highlights the revolutionary potential of deep learning in synthetic chemistry, emphasising how effective and accurate it is in reaction prediction when compared to conventional techniques. AI-driven models in materials science and pharmaceuticals enable the investigation of intricate reaction pathways and the optimisation of synthetic routes by leveraging large datasets from sources such as PubChem and Reaxys. Even with these developments, there are still important problems to be solved, like data accessibility, model interpretability, and integrating AI with experimental techniques. The future of chemical research will be shaped by developments like hybrid models, which integrate AI with conventional methods, and reinforcement learning, in addition to ethical issues. All things considered, deep learning presents previously unheard-of chances to improve reaction prediction and discovery in synthetic chemistry, highlighting the significance of cooperation between chemists, data scientists, and ethicists to guarantee ethical and efficient use of AI technology.

ACKNOWLEDGEMENT:

My sincere appreciation goes out to all the people and institutions whose assistance enabled this review. I am especially grateful to Guru Nanak Dev University, Amritsar for their support and resources during my research. Additionally, I am grateful to the peer reviewers for their insightful comments, which significantly improved the calibre of this paper.

AUTHOR CONTRIBUTION:

Gurjant Singh drafted the manuscript. Gurjant Singh wrote the paper and approved the final manuscript. The author confirms that no paper mill and artifcial intelligence was used.

Data availability Not applicable.

DECLARATIONS:

Ethical approval Not applicable.

Competing interests The authors declare no competing interests

REFERENCES:

1. Campos KR, Coleman PJ, Alvarez JC, Dreher SD, Garbaccio RM, Terrett NK, Tillyer RD, Truppo MD, Parmee ER. The importance of synthetic chemistry in the pharmaceutical industry. Science. 2019 Jan 18;363(6424):eaat0805. Doi: 10.1126/science.aat0805. PMID: 30655413.
2. Niazi SK, Mariam Z. Recent Advances in Machine-Learning-Based Chemoinformatics: A Comprehensive Review. Int J Mol Sci. 2023 Jul 15;24(14):11488. Doi: 10.3390/ijms241411488. PMID: 37511247; PMCID: PMC10380192.
3. Valavanidis, Athanasios. (2023). Artificial Intelligence (AI) is Transforming Most Scientific Areas in Chemistry. Machine-learning tools revolutionize synthesis of fine chemicals, antibiotics and drugs. 1. 1-42.
4. Shoichi Ishida, Kei Terayama, Ryosuke Kojima, Kiyosei Takasu, and Yasushi Okuno. Journal of Chemical Information and Modeling 2022 62 (6), 1357-1367. DOI: 10.1021/acs.jcim.1c01074
5. Chengchun Liu, Yuntian Chen2 and Fanyang Mo. (2024). Transforming organic chemistry research paradigms: Moving from manual efforts to the intersection of automation and artificial intelligence. National Science Open. Volume 3, Number 2. Doi: https://doi.org/10.1360/nso/20230037
6. Ali, Rizvi & Meng, Jiaolong & Khan, Muhammad Ehtisham & Jiang, Xuefeng. (2024). Machine Learning Advancements in Organic Synthesis: A Focused Exploration of Artificial Intelligence Applications in Chemistry. Artificial Intelligence Chemistry. 10.1016/j.aichem.2024.100049.
7. Han R, Yoon H, Kim G, Lee H, Lee Y. Revolutionizing Medicinal Chemistry: The Application of Artificial Intelligence (AI) in Early Drug Discovery. Pharmaceuticals (Basel). 2023 Sep 6;16(9):1259. Doi: 10.3390/ph16091259. PMID: 37765069; PMCID: PMC10537003.
8. Kovács, D.P., McCorkindale, W. & Lee, A.A. Quantitative interpretation explains machine learning models for chemical reaction prediction and uncovers bias. Nat Commun 12, 1695 (2021). https://doi.org/10.1038/s41467-021-21895-w
9. Journals, Stm & Manna, Anindya & Roy, Antara & Rahaman, Rajjakfur & Maiti, Dilip. (2024). Cheminformatics Recent Applications and Influences of Artificial Intelligence (AI) In Chemical and Allied Sciences. 1. 2023. 10.37591/IJCI.
10. Lu, Jieyu & Zhang, Yingkai. (2022). Unified Deep Learning Model for Multitask Reaction Predictions with Explanation. Journal of Chemical Information and Modeling. 62. 10.1021/acs.jcim.1c01467.
11. Chen, LY., Li, YP. Enhancing chemical synthesis: a two-stage deep neural network for predicting feasible reaction conditions. J Cheminform 16, 11 (2024). https://doi.org/10.1186/s13321-024-00805-4
12. Tran TTV, Tayara H, Chong KT. Artificial Intelligence in Drug Metabolism and Excretion Prediction: Recent Advances, Challenges, and Future Perspectives. Pharmaceutics. 2023 Apr 17;15(4):1260. Doi: 10.3390/pharmaceutics15041260. PMID: 37111744; PMCID: PMC10143484.
13. Tu Z, Stuyver T, Coley CW. Predictive chemistry: machine learning for reaction deployment, reaction development, and reaction discovery. Chem Sci. 2022 Nov 28;14(2):226-244. Doi: 10.1039/d2sc05089g. PMID: 36743887; PMCID: PMC9811563.
14. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M, Zhao S. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019 Jun;18(6):463-477. Doi: 10.1038/s41573-019-0024-5. PMID: 30976107; PMCID: PMC6552674.
15. Venkatasubramanian, Venkat & Mann, Vipul. (2021). Artificial intelligence in reaction prediction and chemical synthesis. Current Opinion in Chemical Engineering. 36. 10.1016/j.coche.2021.100749.
16. Ross, John. (2008). Determination of Complex Reaction Mechanisms. Analysis of Chemical, Biological and Genetic Networks †. The journal of physical chemistry. A. 112. 2134-43. 10.1021/jp711313e.
17. Tu Z, Stuyver T, Coley CW. Predictive chemistry: machine learning for reaction deployment, reaction development, and reaction discovery. Chem Sci. 2022 Nov 28;14(2):226-244. Doi: 10.1039/d2sc05089g. PMID: 36743887; PMCID: PMC9811563.
18. Cova TFGG, Pais AACC. Deep Learning for Deep Chemistry: Optimizing the Prediction of Chemical Patterns. Front Chem. 2019 Nov 26;7:809. Doi: 10.3389/fchem.2019.00809. PMID: 32039134; PMCID: PMC6988795.
19. Hirohara, M., Saito, Y., Koda, Y. et al. Convolutional neural network based on SMILES representation of compounds for detecting chemical motif. BMC Bioinformatics 19 (Suppl 19), 526 (2018). https://doi.org/10.1186/s12859-018-2523-5
20. Francesca Grisoni, Michael Moret, Robin Lingwood, and Gisbert Schneider. Journal of Chemical Information and Modeling 2020 60 (3), 1175-1183. DOI: 10.1021/acs.jcim.9b00943
21. Besharatifard, M., Vafaee, F. A review on graph neural networks for predicting synergistic drug combinations. Artif Intell Rev 57, 49 (2024). https://doi.org/10.1007/s10462-023-10669-z
22. Wegner, Joerg & Sterling, Aaron & Guha, Rajarshi & Bender, Andreas & Faulon, Jean-Loup & Hastings, Janna & O'Boyle, Noel & Overington, John & Vlijmen, Herman & Willighagen, Egon. (2012). Cheminformatics. Communications of the ACM. 55. 65-75. 10.1145/2366316.2366334.
23. Jiang, D., Wu, Z., Hsieh, CY. Et al. Could graph neural networks learn better molecular representation for drug discovery? A comparison study of descriptor-based and graph-based models. J Cheminform 13, 12 (2021). https://doi.org/10.1186/s13321-020-00479-8
24. Zhang, C., Zhai, Y., Gong, Z. et al. Transfer learning across different chemical domains: virtual screening of organic materials with deep learning models pretrained on small molecule and chemical reaction data. J Cheminform 16, 89 (2024). https://doi.org/10.1186/s13321-024-00886-1

25. Lawson, Alexander & Swienty-Busch, Jürgen & Geoui, Thibault & Evans, David. (2014). The Making of Reaxys—Towards Unobstructed Access to Relevant Chemistry Information. 10.1021/bk-2014-1164.ch008.
26. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. PubChem Substance and Compound databases. Nucleic Acids Res. 2016 Jan 4;44(D1):D1202-13. Doi: 10.1093/nar/gkv951. Epub 2015 Sep 22. PMID: 26400175; PMCID: PMC4702940.
27. Zheng, Shuangjia & Rao, Jiahua & Zhang, Zhongyue & Xu, Jun & Yang, Yuedong. (2019). Predicting Retrosynthetic Reactions Using Self-Corrected Transformer Neural Networks. Journal of chemical information and modeling. 60. 10.1021/acs.jcim.9b00949.
28. Ginsburg, Boris & Tropsha, Alexander & Isayev, Olexandr. (2020). OpenChem: A Deep Learning Toolkit for Computational Chemistry and Drug Design. 10.26434/chemrxiv.12691943.
29. Struble TJ, Alvarez JC, Brown SP, Chytil M, Cisar J, DesJarlais RL, Engkvist O, Frank SA, Greve DR, Griffin DJ, Hou X, Johannes JW, Kreatsoulas C, Lahue B, Mathea M, Mogk G, Nicolaou CA, Palmer AD, Price DJ, Robinson RI, Salentin S, Xing L, Jaakkola T, Green WH, Barzilay R, Coley CW, Jensen KF. Current and Future Roles of Artificial Intelligence in Medicinal Chemistry Synthesis. J Med Chem. 2020 Aug 27;63(16):8667-8682. Doi: 10.1021/acs.jmedchem.9b02120. Epub 2020 Apr 14. PMID: 32243158; PMCID: PMC7457232.
30. Li, B., Su, S., Zhu, C. et al. A deep learning framework for accurate reaction prediction and its application on high-throughput experimentation data. J Cheminform 15, 72 (2023). https://doi.org/10.1186/s13321-023-00732-w
31. Yongjun Xu, Xin Liu, Xin Cao, Changping Huang, Enke Liu, Sen Qian, Xingchen Liu, Yanjun Wu, Fengliang Dong, Cheng-Wei Qiu, Junjun Qiu, Keqin Hua, Wentao Su, Jian Wu, Huiyu Xu, Yong Han, Chenguang Fu, Zhigang Yin, Miao Liu, Ronald Roepman, Sabine Dietmann, Marko Virta, Fredrick Kengara, Ze Zhang, Lifu Zhang, Taolan Zhao, Ji Dai, Jialiang Yang, Liang Lan, Ming Luo, Zhaofeng Liu, Tao An, Bin Zhang, Xiao He, Shan Cong, Xiaohong Liu, Wei Zhang, James P. Lewis, James M. Tiedje, Qi Wang, Zhulin An, Fei Wang, Libo Zhang, Tao Huang, Chuan Lu, Zhipeng Cai, Fang Wang, Jiabao Zhang, Artificial intelligence: A powerful paradigm for scientific research, The Innovation, Volume 2, Issue 4, 2021, 100179, ISSN 2666-6758, https://doi.org/10.1016/j.xinn.2021.100179.
32. Chen, LY., Li, YP. AutoTemplate: enhancing chemical reaction datasets for machine learning applications in organic chemistry. J Cheminform 16, 74 (2024). https://doi.org/10.1186/s13321-024-00869-2
33. Wu JN, Wang T, Chen Y, Tang LJ, Wu HL, Yu RQ. T-SMILES: a fragment-based molecular representation framework for de novo ligand design. Nat Commun. 2024 Jun 11;15(1):4993. Doi: 10.1038/s41467-024-49388-6. PMID: 38862578; PMCID: PMC11167009.
34. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., … & Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine Learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (pp. 265-283). Retrieved from https://www.tensorflow.org/
35. McKinney, W. (2010). Data Analysis in Python. In Proceedings of the 9th Python in Science Conference (pp. 51-56). Retrieved from https://pandas.pydata.org/
36. Oliphant, T. E. (2006). A Guide to NumPy. Trelgol Publishing. Retrieved from https://numpy.org/
37. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … & VanderPlas, J. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830. Retrieved from https://scikit-learn.org/
38. Rafael, Gozalbes & Barbosa, Frédérique & Nicolaï, Eric & Horvath, Dragos & Froloff, Nicolas. (2009). Development and Validation of a Pharmacophore-Based QSAR Model for the Prediction of CNS Activity. ChemMedChem. 4. 204-9. 10.1002/cmdc.200800282.
39. Borrelli, William & Schrier, Joshua. (2021). Evaluating the Performance of a Transformer-based Organic Reaction Prediction Model. 10.33774/chemrxiv-2021-3nqv9.
40. Pettit RW, Fullem R, Cheng C, Amos CI. Artificial intelligence, machine learning, and deep learning for clinical outcome prediction. Emerg Top Life Sci. 2021 Dec 20;5(6):729–45. Doi: 10.1042/ETLS20210246. Epub ahead of print. PMID: 34927670; PMCID: PMC8786279.
41. Li, B., Su, S., Zhu, C. et al. A deep learning framework for accurate reaction prediction and its application on high-throughput experimentation data. J Cheminform 15, 72 (2023). https://doi.org/10.1186/s13321-023-00732-w
42. Ali, Rizvi & Meng, Jiaolong & Khan, Muhammad Ehtisham & Jiang, Xuefeng. (2024). Machine Learning Advancements in Organic Synthesis: A Focused Exploration of Artificial Intelligence Applications in Chemistry. Artificial Intelligence Chemistry. 10.1016/j.aichem.2024.100049.
43. Ali, K.A., Mohin, S., Mondal, P. et al. Influence of artificial intelligence in modern pharmaceutical formulation and drug development. Futur J Pharm Sci 10, 53 (2024). https://doi.org/10.1186/s43094-024-00625-1
44. Champa-Bujaico E, García-Díaz P, Díez-Pascual AM. Machine Learning for Property Prediction and Optimization of Polymeric Nanocomposites: A State-of-the-Art. Int J Mol Sci. 2022 Sep 14;23(18):10712. Doi: 10.3390/ijms231810712. PMID: 36142623; PMCID: PMC9505448.
45. Strieth-Kalthoff, Felix & Sandfort, Frederik & Kühnemund, Marius & Schäfer, Felix & Kuchen, Herbert & Glorius, Frank. (2022). Machine Learning for Chemical Reactivity The Importance of Failed Experiments. Angewandte Chemie. 134. 10.1002/ange.202204647.
46. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021 Jan;26(1):80-93. Doi: 10.1016/j.drudis.2020.10.010. Epub 2020 Oct 21. PMID: 33099022; PMCID: PMC7577280.

47. Mochen, Liao, et al., (2021). Sustainability implications of artificial intelligence in the chemical industry: A conceptual framework. Journal of Industrial Ecology. Volume 26, Issue 1/P. 164-182. Doi: https://doi.org/10.1111/jiec.13214
48. Carvalho, Diogo V., Eduardo M. Pereira, and Jaime S. Cardoso. 2019. "Machine Learning Interpretability: A Survey on Methods and Metrics" Electronics 8, no. 8: 832. https://doi.org/10.3390/electronics8080832
49. Shah, Varun & Konda, Sreedhar. (2021). Neural Networks and Explainable AI: Bridging the Gap between Models and Interpretability. 5. 163-176. 10.5281/zenodo.10779335.
50. Singh, Shashank & Dhirendra, Pratap & Chandra, Mr. (2024). Enhancing Transparency and Interpretability in Deep Learning Models: A Comprehensive Study on Explainable AI Techniques. 10.55041/IJSREM28675.
51. Tetko IV, Engkvist O, Koch U, Reymond JL, Chen H. BIGCHEM: Challenges and Opportunities for Big Data Analysis in Chemistry. Mol Inform. 2016 Dec;35(11-12):615-621. Doi: 10.1002/minf.201600073. Epub 2016 Jul 28. PMID: 27464907; PMCID: PMC5129546.
52. Shoichi Ishida, Kei Terayama, Ryosuke Kojima, Kiyosei Takasu, and Yasushi Okuno. Journal of Chemical Information and Modeling 2022 62 (6), 1357-1367. DOI: 10.1021/acs.jcim.1c01074
53. Han R, Yoon H, Kim G, Lee H, Lee Y. Revolutionizing Medicinal Chemistry: The Application of Artificial Intelligence (AI) in Early Drug Discovery. Pharmaceuticals (Basel). 2023 Sep 6;16(9):1259. Doi: 10.3390/ph16091259. PMID: 37765069; PMCID: PMC10537003.
54. Zhenpeng Zhou, Xiaocheng Li, and Richard N. Zare. Optimizing Chemical Reactions with Deep Reinforcement Learning. ACS Central Science 2017 3 (12), 1337-1344. DOI: 10.1021/acscentsci.7b00492
55. Zhang C, Zhai Y, Gong Z, Duan H, She YB, Yang YF, Su A. Transfer learning across different chemical domains: virtual screening of organic materials with deep learning models pretrained on small molecule and chemical reaction data. J Cheminform. 2024 Jul 30;16(1):89. Doi: 10.1186/s13321-024-00886-1. PMID: 39080777; PMCID: PMC11290278.
56. Bengani, Vedika. (2024). Hybrid Learning Systems: Integrating Traditional Machine Learning with Deep learning Techniques.. 10.13140/RG.2.2.10461.22244/1.
57. Visan AI, Negut I. Integrating Artificial Intelligence for Drug Discovery in the Context of Revolutionizing Drug Delivery. Life (Basel). 2024 Feb 7;14(2):233. Doi: 10.3390/life14020233. PMID: 38398742; PMCID: PMC10890405.
58. Tran TTV, Tayara H, Chong KT. Artificial Intelligence in Drug Metabolism and Excretion Prediction: Recent Advances, Challenges, and Future Perspectives. Pharmaceutics. 2023 Apr 17;15(4):1260. Doi: 10.3390/pharmaceutics15041260. PMID: 37111744; PMCID: PMC10143484.
59. Hermann, E., Hermann, G. & Tremblay, JC. Ethical Artificial Intelligence in Chemical Research and Development: A Dual Advantage for Sustainability. Sci Eng Ethics 27, 45 (2021). https://doi.org/10.1007/s11948-021-00325-6
60. Elendu C, Amaechi DC, Elendu TC, Jingwa KA, Okoye OK, John Okah M, Ladele JA, Farah AH, Alimi HA. Ethical implications of AI and robotics in healthcare: A review. Medicine (Baltimore). 2023 Dec 15;102(50):e36671. Doi: 10.1097/MD.0000000000036671. PMID: 38115340; PMCID: PMC10727550.

AUTHORS BIOGRAPHY:

Author Name: Gurjant Singh

MSc. Botany (Post Graduate Researcher)

Guru Nanak Dev University, Amritsar

Mr. Gurjant Singh is a Post Graduate Researcher of Guru Nanak Dev University, Amritsar specializing in Biological Sciences and AI in Agriculture Specialist. With over 6 years of experience in the field, Mr. Gurjant Singh has authored number of peer-reviewed papers and contributed to significant advancements in Scientific Academia, particularly in Plant, Artificial Intelligence, Pharmaceutical & Food & Nutrition applications. Mr. Gurjant Singh research focuses on integrating AI to optimize chemical reactions, accelerating drug development processes. He holds a MSc. in Botany from Guru Nanak Dev University, Amritsar.