ISSN 2224-087X. Електроніка та інформаційні технології. 2024. Випуск 27. С. 102–117
Electronics and information technologies. 2024. Issue 27. P. 102–117

UDC 004.8. DOI: https://doi.org/10.30970/eli.27.8

OVERVIEW OF MACHINE LEARNING METHODS FOR ACADEMIC SCHEDULING

O. Zanevych, V. Kukharskyy

Applied Mathematics Department,
Ivan Franko National University of Lviv,
1 Universitetska St., UA–79090 Lviv, Ukraine
oleh.zanevych@lnu.edu.ua,
vitaliy.kukharskyy@lnu.edu.ua

Academic scheduling is assigning some educational activities to available resources, which may include time slots, classrooms, or instructors. Scheduling in the educational field is considerably complicated because of its complexity and dynamic constraints bound to change quickly. Machine learning thus could be a promising solution to data-driven techniques for this complex problem. This article surveys several ML techniques that have been utilized for academic scheduling, such as supervised learning, unsupervised learning, reinforcement learning, and hybrid approaches. We consider the implementation of these methods in increasing the effectiveness and elasticity of the scheduling systems, addressing the particular constraints encountered, and effecting an overall improvement of satisfaction for the students, teachers, and administrators. This will compare, through strengths and weaknesses, different ML techniques to provide insights for the most effective strategies to develop an academic scheduling solution.

Keywords: Academic Scheduling, Machine Learning, Resource Optimization, Educational Institutions, Dynamic Scheduling, Data-Driven Techniques.

Introduction

Efficient academic scheduling forms the base for educational institutes to run smoothly. It means the arranging of classes and examinations along with other academic activities feasibly and possibly by way of the best use of available resources without any conflicting time slot. Effective class schedules support the enhancement of student academics and increased faculty and staff productivity [1]. With curricula becoming more complex and diverse in student needs, traditional scheduling methods are, in most cases, insufficient. The integration of machine learning techniques into academic scheduling processes has, therefore, shown much promise to bring forth many much-needed reforms in how institutions can best plan and manage their resources to meet the diverse requirements of the stakeholders efficiently [2].

Efficient academic scheduling is of importance; however, it is not merely a case of logistical convenience. It affects the overall learning experience, which includes student satisfaction, faculty workloads, and the institutional reputation [3, 4]. A good schedule may result in improved academic performance of students, as they are more inclined to attend classes that have been scheduled at an ideal time for them to do so based on personal and academic commitments. Such scheduling also gives the faculty members an edge by aligning their preferences and availability in a schedule, increasing job satisfaction and teaching quality.

Further, the administration benefits from a more efficient operation with improved facility utilization.

Academic scheduling is full of challenges and multifarious. A significant number of restrictions and variables taken by the significant complications of its task are the availability of faculty members, allocation of rooms, and students' preferences. In academic settings, these factors are further complicated because changes in course offering, enrollment, and staff deployment are not uncommon; hence, flexibility and responsiveness become paramount. Traditional scheduling methods like manual ones or rule-based methods often fail to adjust these complexities, resulting in suboptimal schedules that affect the education experience [5].

One of the most significant academic scheduling challenges is balancing course requirements with faculty availability. Faculty members have different teaching loads, research commitments, and personal likes, which determine their availability time for work. Doing this is a very, very difficult task: making sure each class gets a qualified lecturer at the time it is supposed to be held. Room allocation is also a big headache; different courses have their specific requirements for the size of classrooms and the equipment and location of these classrooms. The whole puzzle of efficient use of space, avoiding conflicts regarding room allocation, yet locating the suitable class for teaching is complex. This further complicates the process. Each student has a course preference, which can engage them with part-time jobs, extracurricular activities, or personal commitments. An optimal schedule should take into consideration such preferences as well to increase the satisfaction and attendance of students. The task is further complex because of institutional policies on limitations for class sizes and prerequisites that must be adhered to in the scheduling.

This paper gives a complete idea of different machine-learning techniques applied in academic scheduling. We would show how to use supervised learning, unsupervised learning, reinforcement learning, and hybrid approaches to solve efficient problems related to the schedule in educational institutions. We attempt to give insight into the strengths and weaknesses of each ML method and use this to guide the design of more effective and flexible solutions for academic scheduling.

Machine learning techniques offer a data-driven approach to scheduling, enabling the analysis and optimization of complex constraints and variables. Supervised learning methods like regression and classification permit the prediction of scheduling outcomes based on historical data. The unsupervised learning techniques, especially clustering and association, would efficiently draw patterns and relationships inside the scheduling data that are hard to see. Here, the reinforcement learning approaches are applied to make dynamic scheduling decisions with real-time adaption of the approaches. Hybrid approaches combine both methods, offering a robust solution for solving the multifaceted challenges of academic scheduling. This article aims to contribute to the ongoing discussion in academic scheduling by presenting an overview of such ML methods. It gives practical insights for a researcher, practitioner, or educational administrator. The practical significance of each method is considered about some success stories in their implementation. We also point out areas in which future research is promising. Through this comprehensive survey, we try to demonstrate the enormous transformational power that machine learning has in enhancing the efficiency, flexibility, and overall quality of the academic scheduling system.

Literature Review

Overview of Academic Scheduling. Academic scheduling refers to the arrangement and provision of events associated with academics, including classes, examinations, and meetings, within an institution. The scheduling process includes the allotment of time and room resources as well as teachers and other necessary facilities for various academic activities in a way that takes place efficiently and effectively. The derived schedule should be student-, faculty-, and administration-friendly and adhere to institutional policies and constraints [1,6].

Challenges. Academic scheduling should, therefore, consider the limitation of resources, such as physical spaces for classes or laboratories and human resources like faculty members. One has to ensure that any class gets an appropriate room for its occupancy and an available instructor at that specific time [7]. This is a huge challenge. Secondly, it is also difficult to balance the various and often conflicting requirements of students, faculty, and policies set by the administration. For example, a student's course preference may conflict with faculty availability or room capacities. It deals with the vast amount of data such as the course catalogs, faculty schedules, room availability, and student enrollments. Managing this data and further processing it to generate an appropriate schedule is quite a challenge. Academic institutions are very dynamic; course offerings, numbers of students enrolled, and faculty being used may change frequently. This all goes on to add a further layer of complexity for schedulers to be flexible and responsive to these changes. Moreover, the schedulers have to make sure that the schedules generated obey the constraints imposed by the institution. Constraints include maximum class size, prerequisite courses, and particular time-slot allocations, among many others [8].

Manual Scheduling. Academic schedules are developed by human schedulers, using expertise and experience. This would typically be done through a gathering of course-offering data, faculty availability, and room capacities. Then, by balancing the various constraints and requirements, time slots and resources are manually allocated to the classes. However, manual scheduling is time-consuming and error-prone. The complexity of managing large datasets and conflicting requirements often leads to suboptimal schedules. Additionally, manual scheduling lacks flexibility and adaptability, making it difficult to respond to changes in real-time.

Heuristics and Rule-Based Systems. Heuristics and rule-based systems are used to improve upon manual scheduling. These methods generate schedules based on predefined rules and heuristics. One such example might be assigning the core courses in a curriculum to large classrooms or preventing any instructor from teaching in two buildings back-to-back during the same day. Although approaches such as this may automate specific segments of the scheduling process and contribute in some way to lightening the load on human schedulers, they do so under heavy compunction. Heuristic and rule-based systems are only as good and complete as the rules they adhere to. Secondly, these systems do not usually adjust to the dynamic and complexity of academic environments; therefore, they may be incapable of creating optimal schedules. Furthermore, they lack the capability of learning from previous experiences to adapt to new situations [9, 10].

Machine learning (ML) is a subset of artificial intelligence that focuses on developing algorithms that enable computers to learn from and make predictions based on data. Unlike traditional programming, where a computer follows explicit instructions, ML algorithms identify patterns and relationships in data to make decisions or predictions. These algorithms can be categorized into several types, including supervised learning, unsupervised learning, reinforcement learning, and hybrid approaches. Each type has its unique characteristics and applications, making ML a versatile tool for solving complex problems across various domains, including academic scheduling [11].

In the context of academic scheduling, ML techniques can be employed to analyze historical scheduling data, uncover hidden patterns, and generate schedules that optimize the use of resources while meeting diverse constraints. By leveraging the vast amounts of data generated by academic institutions, ML models can provide insights that traditional methods might miss. This data-driven approach can significantly enhance the efficiency and effectiveness of the scheduling process, leading to better outcomes for students, faculty, and administrators.

The integration of ML into academic scheduling offers several advantages over traditional methods, making it a compelling solution for modern educational institutions. Here are some of the key benefits:

ML algorithms can process large datasets quickly and accurately, identifying patterns and relationships that would be challenging for human schedulers or rule-based systems to discern. This capability allows for the generation of schedules in a fraction of the time required by manual methods. For instance, ML models can analyze historical data on course enrollments, classroom usage, and faculty availability to predict future scheduling needs and allocate resources accordingly. This efficiency can free up valuable time for administrators, allowing them to focus on other critical tasks [12].

One of the primary strengths of ML is its ability to learn from data and improve over time. As ML models are exposed to more data, they can refine their predictions and generate more accurate schedules. This learning process helps minimize errors and reduce conflicts in the scheduling process. For example, an ML model can learn to avoid scheduling conflicts by recognizing patterns in past data where conflicts occurred and adjusting future schedules accordingly. This accuracy ensures that schedules are not only feasible but also optimized to meet the needs of all stakeholders.

Academic environments are inherently dynamic, with frequent changes in course offerings, enrollment numbers, and faculty assignments [12]. Traditional scheduling methods often struggle to adapt to these changes in real-time. In contrast, ML models can be trained to respond to new data and adjust schedules accordingly. This adaptability is particularly valuable in managing last-minute changes, such as unexpected faculty absences or sudden increases in student enrollment. By continuously learning and adapting, ML-based scheduling systems can maintain optimal schedules even in the face of uncertainty and change.

Academic scheduling involves balancing numerous constraints, including room capacities, faculty availability, student preferences, and institutional policies. These constraints are often interdependent and can conflict with one another, making the scheduling process highly complex. ML algorithms excel at handling such complexity by using advanced techniques to model and optimize multiple constraints simultaneously [12]. For example, reinforcement learning, a type of ML, can be used to develop scheduling policies that maximize resource utilization while minimizing conflicts and meeting all specified constraints. This capability allows for the creation of schedules that are both practical and efficient.

By providing data-driven insights, ML can support better decision-making in the scheduling process. For instance, ML models can analyze trends in course enrollments and predict which courses are likely to have high demand. This information can help administrators make informed decisions about course offerings, faculty assignments, and resource allocation. Additionally, ML can identify potential scheduling issues before they arise, allowing for proactive adjustments and reducing the likelihood of conflicts.

Ultimately, the goal of academic scheduling is to meet the needs of students, faculty, and administrators. ML-based scheduling systems can enhance stakeholder satisfaction by creating schedules that consider individual preferences and constraints. For students, this means having access to courses that fit their academic plans and personal commitments [11]. For faculty, it means teaching schedules that align with their availability and preferences. For administrators, it means streamlined operations and efficient use of resources. By optimizing schedules for all stakeholders, ML can contribute to a more positive and productive academic environment.

Numerous case studies and real-world applications demonstrate the effectiveness of ML in academic scheduling. For example, some universities have implemented ML-based systems to optimize classroom assignments, resulting in significant improvements in room utilization and scheduling efficiency. Other institutions have used ML to develop personalized student schedules, enhancing student satisfaction and academic performance. These successes highlight the transformative potential of ML in academic scheduling and underscore its relevance in addressing the challenges faced by educational institutions.

The field of ML in academic scheduling is continuously evolving, with ongoing research exploring new methods and applications. Future developments may include the integration of more advanced ML techniques, such as deep learning, to further enhance scheduling accuracy and adaptability. Additionally, the incorporation of real-time data from sources like campus management systems and student information systems could enable even more responsive and dynamic scheduling solutions. As ML technology advances, its application in academic scheduling is likely to become increasingly sophisticated, offering even greater benefits to educational institutions.

Materials and Methods

Supervised Learning. Supervised learning is a type of machine learning where the model is trained on labeled data, meaning that each training example is paired with an output label. The goal is for the model to learn the relationship between inputs and outputs, allowing it to make accurate predictions on new, unseen data. In the context of academic scheduling, supervised learning techniques can be employed to predict resource requirements, classify courses, and optimize scheduling decisions (See Table 1).

Linear Regression. Linear regression is a fundamental supervised learning algorithm used for predicting a continuous output variable based on one or more input features. In academic scheduling, linear regression can be utilized to predict resource requirements such as the number of classrooms needed, the demand for specific courses, or the availability of faculty members [12].

For example, by analyzing historical data on course enrollments and classroom utilization, a linear regression model can predict the number of students likely to enroll in a particular course in the upcoming semester. This prediction can then inform decisions about how many sections of the course to offer and what size classrooms will be needed. Linear regression helps in understanding trends and making data-driven decisions to ensure that resources are allocated efficiently.

Decision Trees and Random Forests. Decision trees are a type of supervised learning algorithm used for both classification and regression tasks. They work by splitting the data into subsets based on the value of input features, creating a tree-like structure where each node represents a decision based on a feature, and each branch represents the outcome of that decision [13].

In academic scheduling, decision trees can be applied to classification problems, such as categorizing courses by type (e.g., lecture, lab, seminar) and scheduling them accordingly. For instance, a decision tree model can classify courses based on their characteristics (e.g., duration, required equipment, expected enrollment) and assign them to appropriate time slots and rooms.

Random forests, an ensemble method that combines multiple decision trees, can enhance the accuracy and robustness of these classifications. By aggregating the predictions of many trees, random forests reduce the risk of overfitting and improve generalization to new data. This approach can be particularly useful in handling the complex and variable nature of academic scheduling.

Table 1

Method	Description	Applications in Academic Scheduling	Advantages	Limitations
Linear Regression	Predicts a continuous output variable based on input features	Predicting resource requirements such as classroom needs, course enrollments	Simple to implement, interpretable results	Assumes a linear relationship, sensitive to outliers
Decision Trees	Splits data into subsets based on feature values, creating a tree structure	Classifying courses by type, assigning time slots and rooms	Easy to visualize, handles categorical and numerical data	Prone to overfitting, can be unstable with small changes in data
Random Forests	Ensemble method combining multiple decision trees	Enhancing classification accuracy, reducing overfitting	Reduces overfitting, handles large datasets well	Can be computationally intensive, less interpretable than single decision trees
Support Vector Machines	Finds the hyperplane that best separates data points of different classes	Classifying courses, predicting optimal number of sections	Effective in high-dimensional spaces, robust to overfitting	Can be computationally intensive, sensitive to choice of kernel and parameters

Support Vector Machines (SVM). Support Vector Machines are supervised learning algorithms that can be used for both classification and regression tasks. SVMs work by finding the hyperplane that best separates data points of different classes in a high-dimensional space. For classification tasks, SVMs aim to maximize the margin between different classes, while for regression tasks, they aim to fit the best hyperplane to the data [14].

In the context of academic scheduling, SVMs can be employed to classify courses and predict scheduling outcomes based on multiple features. For example, an SVM model can classify courses into categories such as high-demand or low-demand, based on features like past enrollment numbers, course difficulty, and student feedback. This classification can help in prioritizing the scheduling of high-demand courses to ensure they are offered at convenient times and locations.

ISSN 2224-087X. Electronics and information technologies. 2024. Issue 27

Additionally, SVMs can be used for regression tasks such as predicting the optimal number of sections for a course or estimating the required number of instructors for a department. By accurately predicting these variables, SVM models can contribute to more effective and efficient scheduling decisions.

Implementation and Evaluation. To implement these supervised learning methods, the first step is to collect and preprocess relevant data, including historical scheduling information, course attributes, enrollment figures, and faculty availability. The data is then divided into training and test sets, with the training set used to build the model and the test set used to evaluate its performance.

The performance of the models is assessed using metrics such as accuracy, precision, recall, and mean squared error, depending on whether the task is classification or regression. Cross-validation techniques can be employed to ensure the robustness and generalizability of the models. Once validated, the models can be integrated into the scheduling system to assist in decision-making and optimize the scheduling process.

Unsupervised Learning. Unsupervised learning involves training algorithms on data without labeled outcomes. The goal is to uncover hidden patterns and structures within the data [15,16]. In academic scheduling, unsupervised learning techniques can provide valuable insights by grouping similar courses or schedules and identifying key features that influence scheduling decisions (See Table 2***)***.

Table 2

Method	**Description**	**Applications in Academic Scheduling**	**Advantages**	**Limitations**
Clustering (e.g., K-means)	Groups similar data points into clusters based on their features	Grouping similar courses or schedules, identifying patterns in student schedules	Simple to implement, intuitive results, useful for finding natural groupings	Choosing the number of clusters (K) can be arbitrary, sensitive to initial centroids, may not handle complex data structures well
Principal Component Analysis	Reduces dimensionality of data by transforming it into a set of principal components	Reducing complexity of scheduling data, identifying key features and patterns	Reduces dimensionality, helps in data visualization, highlights most significant features	Loss of information, interpretation of principal components can be challenging, assumes linear relationships

Clustering (e.g., K-means). Clustering is a technique used to group similar data points based on their features. K-means clustering, one of the most popular clustering methods, partitions data into K distinct clusters, where each data point belongs to the cluster with the nearest mean. [15]

In academic scheduling, clustering can be applied to group similar courses or schedules, facilitating more efficient resource allocation. For example, courses can be clustered based on attributes such as enrollment numbers, course type (e.g., lecture, lab, seminar), and required resources (e.g., specialized equipment, room size). By grouping similar courses together, schedulers can identify patterns and make informed decisions about room assignments and time slots. For instance, high-enrollment courses that require large classrooms can be clustered together and scheduled in appropriately sized rooms.

Clustering can also be used to analyze student schedules [15]. By grouping students with similar course selections, institutions can identify common scheduling patterns and potential bottlenecks. This information can help in designing schedules that minimize conflicts and maximize the availability of preferred courses.

Principal Component Analysis (PCA). Principal Component Analysis is a dimensionality reduction technique used to transform high-dimensional data into a lower-dimensional space while retaining most of the original variability. PCA identifies the principal components, which are the directions of maximum variance in the data, and projects the data onto these components [16].

In academic scheduling, PCA can be employed to reduce the dimensionality of complex scheduling data and identify the key features that influence scheduling decisions. For example, scheduling data may include numerous variables such as course attributes, faculty availability, room characteristics, and student preferences. By applying PCA, these high-dimensional data can be transformed into a smaller set of principal components that capture the most significant patterns and relationships.

This reduction in dimensionality simplifies the analysis and visualization of the data, making it easier to identify trends and make data-driven decisions. For instance, PCA can highlight the most important factors affecting course enrollment or room utilization, enabling schedulers to prioritize these factors when creating schedules. Additionally, PCA can help in identifying correlations between different scheduling variables, providing insights into how changes in one aspect of the schedule might impact others.

Implementation and Evaluation. To implement these unsupervised learning methods, the first step is to collect and preprocess relevant scheduling data. This data might include course attributes, enrollment figures, faculty availability, room characteristics, and student preferences. Once the data is prepared, clustering algorithms like *K*-means and dimensionality reduction techniques like PCA can be applied.

The performance of clustering methods can be evaluated using metrics such as the silhouette score, which measures how similar each data point is to its own cluster compared to other clusters. Visualization tools such as scatter plots and cluster heatmaps can also be used to assess the quality and interpretability of the clusters. For PCA, the explained variance ratio is a key metric that indicates how much of the original data's variability is captured by each principal component [16]. Visualizing the principal components can provide insights into the structure and relationships within the data.

By leveraging clustering and PCA, academic institutions can gain a deeper understanding of their scheduling data, uncover hidden patterns, and make more informed decisions that optimize resource utilization and enhance the overall scheduling process.

Reinforcement Learning. Reinforcement learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment to maximize a cumulative reward [17]. In academic scheduling, RL can be particularly effective in optimizing scheduling policies through continuous learning and adaptation (See Table 3).

Table 3

Method	Description	Applications in Academic Scheduling	Advantages	Limitations
Markov Decision Processes	Mathematical framework for modeling decision-making with states, actions, transition probabilities, and rewards	Modeling scheduling configurations, finding optimal scheduling policies	Provides a systematic approach, handles uncertainty, can optimize over long-term	Requires accurate modeling of states and transitions, computationally intensive for large state spaces
Q-Learning	Model-free reinforcement learning algorithm that learns the value of actions in different states through trial and error	Optimizing scheduling policies, dynamically adapting schedules	Does not require a model of the environment, learns directly from interactions, flexible and adaptive	Requires extensive training, can be slow to converge, sensitive to learning parameters

Markov Decision Processes (MDP). Markov Decision Processes (MDPs) are a mathematical framework used to model decision-making problems where outcomes are partly random and partly under the control of a decision-maker. An MDP is defined by a set of states, a set of actions, transition probabilities between states, and rewards associated with state transitions. The objective is to find a policy, a strategy for choosing actions, that maximizes the expected cumulative reward over time [17].

In the context of academic scheduling, MDPs can be used to model the scheduling process where each state represents a specific configuration of the schedule, actions represent possible scheduling decisions (e.g., assigning a course to a particular time slot), and rewards reflect the quality of the scheduling decisions (e.g., minimization of conflicts, maximization of resource utilization).

For example, the states in an MDP for academic scheduling could include various combinations of course assignments, room allocations, and time slots. Actions might involve assigning a course to a specific time slot or swapping the schedules of two courses. The transition probabilities would indicate the likelihood of moving from one scheduling configuration to another based on these actions. Rewards could be designed to penalize conflicts (e.g., double-booking of rooms or instructors) and to reward efficient use of resources and adherence to constraints.

Using MDPs allows for systematic exploration of different scheduling configurations and the development of optimal policies that balance competing constraints and objectives.

Q-Learning. Q-learning is a model-free reinforcement learning algorithm that seeks to learn the value of actions in different states, which is represented by a Q-value. The Q-value of a state-action pair represents the expected cumulative reward of taking that action in that state and following the optimal policy thereafter. The agent updates Q-values through trial and error interactions with the environment, progressively improving its policy [18].

In academic scheduling, Q-learning can be used to optimize scheduling policies by iteratively refining the assignment of courses to time slots, rooms, and instructors. The process involves [18]:

1. ***Initialization:*** Initialize the Q-values for all state-action pairs arbitrarily, or with some heuristic values.
2. ***Interaction:*** For each scheduling iteration, the agent observes the current state of the schedule, selects an action (e.g., assigning a course to a time slot) based on an exploration-exploitation strategy, and executes it.
3. ***Reward Calculation:*** After executing the action, the agent receives a reward based on the quality of the resulting schedule (e.g., how well it meets constraints and optimizes resource usage).
4. ***Q-value Update:*** The Q-value for the state-action pair is updated based on the reward received and the maximum Q-value of the next state. The update rule is [18]:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)] \quad (1)$$

   where s is the current state, a is the action taken, r is the reward received, s' is the next state, a' is the action performed in the next state, α is the learning rate, and γ is the discount factor.
5. ***Policy Improvement:*** Over time, as the Q-values converge, the agent develops an optimal policy for scheduling that maximizes the cumulative reward.

By using Q-learning, academic scheduling systems can dynamically adapt to changes and continuously improve their performance through experience. This method allows for flexible and robust optimization of scheduling policies, even in complex and dynamic academic environments.

Implementation and Evaluation. To implement MDPs and Q-learning in academic scheduling, the first step is to define the states, actions, transition probabilities, and reward structure that accurately reflect the scheduling environment. The RL agent is then trained using historical scheduling data and simulations to iteratively improve its policy.

The performance of the RL-based scheduling system can be evaluated using metrics such as the average reward, the number of scheduling conflicts, resource utilization efficiency, and stakeholder satisfaction. Comparing the RL-based approach with traditional scheduling methods can provide insights into its effectiveness and potential for improvement.

By leveraging reinforcement learning techniques such as MDPs and Q-learning, academic institutions can develop adaptive and optimized scheduling solutions that continuously learn and improve from experience, leading to more efficient and effective resource allocation and scheduling outcomes.

Case Studies and Applications

Real-World Examples. In the academic scheduling domain, the application of machine learning techniques has already yielded auspicious results across different educational institutions. Real-world examples demonstrate the practical benefits gained and improvements made by applying various ML methods. Institutions have also been able to optimize many different ways of scheduling processes, reduce conflicts, and increase resource utilization in many ways by taking the help of historical data along with advanced algorithms. What follows are successful implementation examples of ML methods in academic scheduling.

Example 1: University of Melbourne. The University of Melbourne implemented a machine learning-based scheduling system to optimize course timetabling. The system utilized a combination of supervised learning techniques, including decision trees and support vector machines (SVM), to classify courses and predict resource requirements. This ML-based approach significantly reduced scheduling conflicts and improved the allocation of classrooms and instructors [19].

Example 2: Technical University of Munich. The Technical University of Munich applied reinforcement learning, specifically Q-learning, to optimize their exam scheduling process. The system was designed to dynamically adapt to changes in student enrollments and faculty availability. By continuously learning from past scheduling outcomes, the Q-learning model improved the overall efficiency of the exam scheduling process and minimized conflicts [20].

Example 3: Stanford University. Stanford University employed clustering and principal component analysis (PCA) to analyze and optimize student schedules. Clustering techniques, such as K-means, were used to group students with similar course preferences, while PCA helped in reducing the dimensionality of the scheduling data to identify key factors influencing course selections. This approach enhanced the flexibility and accuracy of student schedule planning, resulting in higher student satisfaction and better resource utilization [21].

Comparative Analysis. To compare the performance of different ML methods based on these case studies, several key metrics can be considered, including scheduling accuracy, conflict reduction, computational efficiency, and adaptability to changes.

This Table 4 provides a concise comparison of the performance of supervised learning, reinforcement learning, and unsupervised learning methods in the context of academic scheduling, highlighting their respective strengths and limitations. By analyzing these examples and metrics, it becomes evident that different ML techniques can offer unique benefits depending on the specific requirements and constraints of the scheduling problem at hand. Combining these methods in a hybrid approach could leverage their respective advantages, resulting in a more comprehensive and effective scheduling system.

Challenges and Future Directions

Current Limitations. While machine learning (ML) methods have shown significant promise in improving academic scheduling, several limitations still need to be addressed to fully realize their potential.

The effectiveness of ML algorithms heavily relies on the quality and quantity of data available for training. In academic scheduling, data quality issues such as missing values, inaccuracies, and inconsistencies can hinder the performance of ML models. For instance, incomplete records of faculty availability or errors in room allocation data can lead to suboptimal scheduling outcomes. Additionally, historical scheduling data may not always reflect the latest changes in academic policies or resource availability, which can affect the relevance and accuracy of the models.

Table 4

Metric	**Supervised Learning (University of Melbourne)**	**Reinforcement Learning (Technical University of Munich)**	**Unsupervised Learning (Stanford University)**
Accuracy and Conflict Reduction	Significant reduction in scheduling conflicts and improved resource allocation	Noticeable decrease in exam scheduling conflicts through continuous learning	Contributed to balanced course distribution and reduced conflict likelihood
Computational Efficiency	Requires substantial computational resources for training; quick scheduling once trained	Requires extensive training and computational power to converge; ongoing learning	Computationally efficient for analyzing and reducing data complexity
Adaptability	Robust initial solutions; less flexible in real-time without retraining	High adaptability through continuous learning and policy updates	Supports flexible decision-making; limited direct real-time adaptability

ML algorithms, particularly those used in supervised and reinforcement learning, often require substantial computational resources for training and optimization. Complex models such as deep neural networks or extensive Q-learning algorithms can be computationally intensive, making them challenging to implement in real-time scheduling systems. The need for high computational power can also limit the accessibility of these advanced techniques for smaller institutions with limited technological infrastructure.

As the size of academic institutions and the complexity of their scheduling requirements increase, the scalability of ML methods becomes a critical concern. Algorithms that perform well on smaller datasets may struggle to maintain efficiency and accuracy when applied to larger and more diverse datasets. Ensuring that ML models can scale effectively to handle the growing demands of large educational institutions remains a significant challenge.

Many advanced ML models, such as deep learning networks, operate as "black boxes," making it difficult to interpret their decision-making processes. In the context of academic scheduling, where transparency and explainability are important for gaining stakeholder trust, the lack of interpretability can be a drawback. Decision-makers may hesitate to adopt ML solutions if they cannot understand or justify the reasoning behind the generated schedules.

Future Research. To overcome these limitations and further enhance the application of ML in academic scheduling, several areas of future research are suggested.

Future research should focus on developing more efficient ML algorithms that can handle the computational complexity and scalability issues associated with large scheduling datasets. Techniques such as approximate algorithms, parallel processing, and optimized hardware implementations (e.g., using GPUs) can help reduce the computational burden. Additionally, exploring hybrid approaches that combine ML with traditional optimization methods could lead to more efficient and effective scheduling solutions.

Academic environments are dynamic, with frequent changes in course offerings, student enrollments, and faculty assignments. Incorporating real-time data and dynamically adapting to these changes is crucial for maintaining optimal schedules. Future research should investigate methods for integrating dynamic data streams into ML models, enabling continuous learning and real-time updates. Techniques such as online learning and adaptive algorithms can help create more responsive scheduling systems.

Improving the quality of input data is essential for the success of ML models in academic scheduling. Future research should focus on developing robust data preprocessing techniques to handle missing values, correct inaccuracies, and ensure consistency. Implementing standardized data collection protocols and leveraging data augmentation methods can also enhance the quality and availability of training data.

To address the interpretability challenge, future research should explore methods for making ML models more transparent and explainable. Techniques such as model simplification, rule extraction, and visualization tools can help stakeholders understand the decision-making processes of ML algorithms. Additionally, developing user-friendly interfaces that present scheduling recommendations along with explanations can facilitate the adoption of ML-based scheduling systems.

The field of machine learning is rapidly evolving, with new techniques and approaches continuously emerging. Future research should stay abreast of these developments and explore their potential applications in academic scheduling. For example, advancements in deep learning, reinforcement learning, and transfer learning could offer novel solutions to existing scheduling challenges. Investigating the use of these cutting-edge techniques can lead to more innovative and effective scheduling systems.

Finally, fostering cross-disciplinary collaboration between computer scientists, educators, and domain experts is crucial for advancing research in ML-based academic scheduling. Collaborative efforts can ensure that the developed models are tailored to the specific needs and constraints of educational institutions, leading to more practical and impactful solutions.

Conclusion

This article has provided a comprehensive overview of various machine learning (ML) methods applied to academic scheduling. We began by highlighting the importance of efficient academic scheduling and the challenges faced by traditional methods, including resource constraints, conflicting requirements, and large datasets. The introduction of supervised learning techniques such as linear regression, decision trees, random forests, and support vector machines (SVM) has shown how these methods can predict resource requirements and optimize scheduling decisions. Unsupervised learning techniques, including clustering and principal component analysis (PCA), have demonstrated their value in grouping similar courses and reducing data complexity.

Reinforcement learning methods, particularly Markov Decision Processes (MDP) and Q-learning, have been discussed for their potential to optimize scheduling policies through continuous learning and adaptation. Real-world examples from institutions like the University of Melbourne, Technical University of Munich, and Stanford University have illustrated the successful implementation of these ML techniques in improving scheduling efficiency and reducing conflicts.

We also addressed the current limitations of ML methods in academic scheduling, such as data quality issues, computational complexity, scalability, and interpretability. Future research directions were suggested, including improving algorithm efficiency, incorporating dynamic data, enhancing data quality, increasing model interpretability, exploring new ML techniques, and fostering cross-disciplinary collaboration.

The potential impact of machine learning on academic scheduling is profound. By leveraging data-driven techniques, educational institutions can significantly enhance the efficiency, accuracy, and adaptability of their scheduling processes. ML methods offer the promise of optimized resource utilization, reduced conflicts, and improved stakeholder satisfaction, ultimately contributing to a more effective and flexible educational environment.

As the field of machine learning continues to evolve, there are vast opportunities for further exploration and innovation in academic scheduling. Future research and development efforts should focus on addressing the current limitations and expanding the capabilities of ML-based scheduling systems. By doing so, educational institutions can stay ahead of the dynamic demands of academic environments and ensure that their scheduling processes are both efficient and responsive.

In conclusion, the integration of machine learning into academic scheduling represents a transformative approach that holds the potential to revolutionize the way educational institutions manage their resources and meet the diverse needs of their stakeholders. Encouraging further exploration and adoption of these advanced techniques will pave the way for more intelligent, adaptive, and efficient scheduling solutions in the future.

References

[1] *Burke, E. K., & Petrovic, S.* (2002). Recent research directions in automated timetabling. *European Journal of Operational Research*, 140(2), 266-280. https://doi.org/10.1016/S0377-2217(02)00069-3

[2] *Sagir, M., & Ozturk, Z. K.* (2010). Exam scheduling: Mathematical modeling and parameter estimation with the Analytic Network Process approach. *Mathematical and Computer Modelling*, *52*(5–6), 930–941. https://doi.org/10.1016/j.mcm.2010.05.029

[3] *Abdipoor, S., Yaakob, R., Goh, S. L., & Abdullah, S.* (2023). Meta-heuristic approaches for the University Course Timetabling Problem. *Intelligent Systems With Applications*, *19*, 200253. https://doi.org/10.1016/j.iswa.2023.200253

[4] *Song, T., Liu, S., Tang, X., Peng, X., & Chen, M.* (2018). An iterated local search algorithm for the University Course Timetabling Problem. *Applied Soft Computing*, *68*, 597–608. https://doi.org/10.1016/j.asoc.2018.04.034

[5] *Algasm, W. A.* (2020). Hybrid Algorithm to Solve Timetabling Problem. *IOP Conference Series. Materials Science and Engineering*, *928*(3), 032053. https://doi.org/10.1088/1757-899x/928/3/032053

[6] *Özcan, E., Bilgin, B., & Korkmaz, E. E.* (2008). A comprehensive analysis of hyper-heuristics. *Intelligent Data Analysis*, *12*(1), 3–23. https://doi.org/10.3233/ida-2008-12102

[7] *De Causmaecker, P., & Vanden Berghe, G.* (2011). A decomposed metaheuristic approach for a real-world university timetabling problem. *European Journal of Operational Research*, 212(3), 530-546. http://dx.doi.org/10.1016/j.ejor.2008.01.043

[8] *Kingston, J. H.* (2011). Hierarchical timetabling. *Annals of Operations Research*, 194(1), 53-64. doi:10.1007/s10479-010-0760-2

[9] *Kaviani, M., Shirouyehzad, H., Sajadi, S. M., & Salehi, M.* (2014). A heuristic algorithm for the university course timetabling problems by considering measure index: a case study. *International Journal of Services and Operations Management*, *18*(1), 1. https://doi.org/10.1504/ijsom.2014.060448

[10] *Rezaeipanah, A., Matoori, S. S., & Ahmadi, G.* (2020). A hybrid algorithm for the university course timetabling problem using the improved parallel genetic algorithm and local search. *Applied Intelligence*, *51*(1), 467–492. https://doi.org/10.1007/s10489-020-01833-x

[11] *Van Den Bergh, J., Beliën, J., De Bruecker, P.,* Demeulemeester, E., & De Boeck, L. (2013). Personnel scheduling: A literature review. *European Journal of Operational Research*, *226*(3), 367–385. https://doi.org/10.1016/j.ejor.2012.11.029

[12] *James, G., Witten, D., Hastie, T., & Tibshirani, R.* (2021). An Introduction to Statistical Learning. In *Springer texts in statistics*. https://doi.org/10.1007/978-1-0716-1418-1

[13] *Fawagreh, K., Gaber, M. M., & Elyan, E.* (2014). Random forests: from early developments to recent advancements. *Systems Science & Control Engineering*, *2*(1), 602–609. https://doi.org/10.1080/21642583.2014.956265

[14] *Hao, P. Y., Chiang, J. H., & Chen, Y. D.* (2022). Possibilistic classification by support vector networks. *Neural Networks*, *149*, 40–56. https://doi.org/10.1016/j.neunet.2022.02.007

[15] *Oyewole, G. J., & Thopil, G. A.* (2022). Data clustering: application and trends. *Artificial Intelligence Review*, *56*(7), 6439–6475. https://doi.org/10.1007/s10462-022-10325-y

[16] *Jolliffe, I. T., & Cadima, J.* (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions - Royal Society. Mathematical, Physical and Engineering Sciences/Philosophical Transactions - Royal Society. Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

[17] *Sutton, R. S., & Barto, A. G.* (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press. Retrieved from http://incompleteideas.net/book/the-book-2nd.html

[18] *Jang, B., Kim, M., Harerimana, G., & Kim, J. W.* (2019). Q-Learning Algorithms: A Comprehensive Classification and Applications. *IEEE Access*, *7*, 133653–133667. https://doi.org/10.1109/access.2019.2941229

[19] Technical University of Munich. (n.d.). Application of Q-learning for exam scheduling optimization. Retrieved from TUM Research Database

[20] University of Melbourne. (n.d.). Machine learning-based course timetabling. Retrieved from University of Melbourne Research Database

[21] Stanford University. (n.d.). Clustering and PCA for student schedule optimization. Retrieved from Stanford University Research Database

ОГЛЯД МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ СКЛАДАННЯ АКАДЕМІЧНОГО РОЗКЛАДУ

О. Заневич, В. Кухарський

Кафедра прикладної математики,
Львівський національний університет імені Івана Франка,
вул. Університетська, 1, 79090, Львів, Україна
oleh.zanevych@lnu.edu.ua,
vitaliy.kukharskyy@lnu.edu.ua

У статті розглядається важлива проблема планування академічних занять в освітніх установах, що вимагає постійного управління великою кількістю динамічних і різноманітних ресурсів. Зосереджується увага на можливостях машинного навчання (МН), які можуть значно оптимізувати процес складання розкладу за допомогою обробки великих обсягів даних і знаходження ефективних рішень, що забезпечують збалансоване використання ресурсів і задоволення потреб всіх зацікавлених сторін.

Основна увага в статті приділяється декільком основним видам МН: контрольоване навчання, неконтрольоване навчання, навчання з підкріпленням та гібридні підходи. Кожен з цих підходів аналізується на предмет його здатності оптимізувати планування академічних занять, враховуючи обмеження і вимоги, які часто змінюються. Наприклад, контрольоване навчання використовується для прогнозування потреб у ресурсах на основі історичних даних, тоді як неконтрольоване навчання може допомогти виявити приховані шаблони і взаємозв'язки без заздалегідь відомих результатів.

Навчання з підкріпленням, зокрема, підходи, що базуються на процесах прийняття рішень Маркова і Q-навчання, представляє собою перспективний напрямок для адаптивного і гнучкого планування, що може оперативно реагувати на зміни в умовах і даних. У статті наводяться приклади успішного використання цих технік у різних університетах, які продемонстрували підвищення ефективності розкладу і зменшення конфліктів у ресурсах.

У статті підкреслюється необхідність подальших досліджень і співпраці між фахівцями для вдосконалення використання МН в академічному плануванні. Інноваційні підходи та більш глибока інтеграція МН можуть трансформувати процеси планування в освіті, підвищуючи ефективність і адаптивність до мінливих умов. Потенціал МН для оптимізації навчальних розкладів у сучасній освіті постійно зростає, що обумовлює потребу в нових подальших дослідженнях. Ця стаття сприяє обговоренню застосування МН в академічному плануванні, висвітлюючи досягнення та проблеми, з якими доводиться стикатися.

Ключові слова: академічний розклад, машинне навчання, оптимізація ресурсів, освітні заклади, динамічне планування, методи, керовані даними.

The article was received by the editorial office on 25.06.2024.
Accepted for publication on 01.07.2024.