



A Hybrid Machine Learning Approach in Educational Data Mining for Academic Risk Classification

Made Pasek Agus Ariawan^{1,*}, Ni Putu Eka Apriyanthi¹, and
Ida Bagus Adisimakrisna¹

¹ Information Technology Department, Politeknik Negeri Bali, Indonesia
pasekagus@pnb.ac.id

Abstract. Higher education institutions currently face major challenges in managing increasingly complex and diverse student academic data. The available data encompasses exam scores, Grade Point Average (GPA), attendance, and learning activities on e-learning platforms. If analyzed correctly, this data can serve as a vital foundation for detecting students at risk of academic failure. However, traditional methods are often less than optimal when dealing with data heterogeneity. K-Nearest Neighbor (KNN) is known to be effective in proximity-based classification, but it has limitations when confronting complex data distributions. On the other hand, K-means clustering can group students based on similar characteristics, yet it does not provide a clear final classification. Consequently, there is a growing need to develop a hybrid approach that combines the strengths of both KNN and K-means, allowing the academic risk classification process to be carried out more accurately and adaptively. This study utilizes an Educational Data Mining (EDM) approach with a hybrid KNN + K-means model. The research stages begin with the collection of student academic data, including exam scores, GPA, attendance, and learning activities. This data is then processed through K-means clustering to discover patterns and groups of students with similar characteristics. The clustering results are subsequently used as the basis for KNN classification, which determines risk categories based on proximity to labeled student data. The resulting model will be evaluated using accuracy, precision, recall, and F1-score metrics to ensure the reliability of the outcomes.

Keywords: Academic, Classification, EDM, Hybrid Machine Learning, KNN

1 Introduction

The development of information technology has brought significant changes to the educational landscape, particularly in the management of student academic data. Higher education institutions now have access to vast amounts of data, ranging from exam scores, attendance, and learning activities to interactions within e-learning platforms. When analyzed appropriately, this abundant data can provide crucial insights for the early detection and classification of students' academic risks. One rapidly growing approach is Educational Data Mining (EDM), which involves applying data mining

techniques to explore hidden patterns within educational data (Saputra et al., 2025). EDM enables educational institutions to identify students who potentially face academic difficulties so that timely interventions such as tutoring, counseling, or remedial programs can be implemented (Farouk et al., 2023).

In the context of academic risk classification, the K-Nearest Neighbor (KNN) method is widely recognized for its effectiveness due to its capability to classify data based on proximity to other data points with known categories (Nadia Bakri & Sofinah Harahap, 2025). However, KNN possesses limitations when dealing with unstructured data or complex data distributions. Conversely, the K-means clustering method can facilitate the grouping of data into specific clusters based on shared characteristics, thereby rendering risk patterns more apparent (Lu et al., 2025)(Ariawan et al., 2023, 2024). By integrating both methods into a hybrid KNN + K-means framework, the academic risk classification process is expected to become more accurate and adaptive. In this architecture, K-means functions to group students based on their academic characteristics, while KNN is utilized to determine the final classification based on proximity to the labeled data. This hybrid approach not only enhances accuracy but also provides a more comprehensive overview of student risk profiles (Patidar et al., 2025) (Guozhang et al., 2025).

Research on academic risk classification using EDM with a hybrid KNN + K-means method is imperative. It assists educational institutions in performing early detection of at-risk students, provides a more objective decision-making basis for designing learning strategies, and supports the creation of a more adaptive education system oriented toward student success (Alfiah et al., 2026). Based on the described background, the primary problem in higher education is formulated as how to detect and classify students at risk of academic failure more accurately and efficiently. Although student academic data is available in large volumes, its utilization remains suboptimal due to the limitations of the analysis methods employed. Traditional classification methods, such as K-Nearest Neighbor (KNN), have proven effective in grouping students based on data proximity, yet they exhibit weaknesses when confronting complex and heterogeneous data. On the other hand, the K-means clustering method is capable of grouping students based on similar characteristics but falls short of providing a definitive final classification. Consequently, there is an emerging need to develop a hybrid KNN + K-means approach that synthesizes the advantages of both methods.

2 Methodology

The study commences with a Preliminary Study phase to strengthen the theoretical foundation, followed by Problem Identification to formulate the primary objectives to be achieved. Once the research focus is established, a Data Collection process is conducted from relevant sources, succeeded by the Data Processing (preprocessing) phase to ensure the raw data is clean and prepared for modeling. This processed data then enters the core modeling phase, which consists of the Clustering Process, followed by the Classification Process, and concludes with the Results Analysis to evaluate the model's performance before the research is officially finalized. The systematic steps involved in this development are illustrated in Figure 1

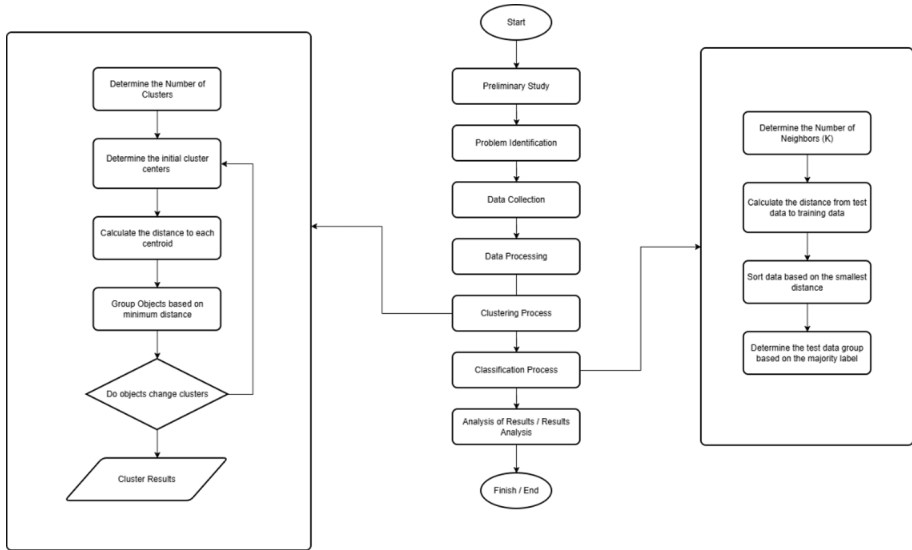


Figure 1. Research Method

This section outlines the primary framework of the research from inception to completion

- 1) **Start:** Initiating the research process.
- 2) **Preliminary Study:** Conducting a literature review, identifying theoretical references, and studying prior research relevant to the selected topic.
- 3) **Problem Identification:** Formulating the problems to be addressed and defining the research objectives.
- 4) **Data Collection:** Retrieving and gathering the required raw data from interconnected data sources.
- 5) **Data Processing:** Performing data preprocessing, which includes cleaning missing values, executing normalization, or applying data transformations to ready the data for modeling.
- 6) **Clustering Process:** Inputting the prepared data into the unlabeled grouping stage (the sub-processes are detailed in the Left Box).
- 7) **Classification Process:** Classifying the clustering results or new data points using a neighbor-proximity-based method (the sub-processes are detailed in the Right Box).
- 8) **Results Analysis:** Evaluating the model's performance, analyzing the formed cluster patterns, and interpreting the classification results.
- 9) **Finish / End:** The research concludes once the objectives have been successfully achieved

Clustering Sub-Process: K-Means (Left Box)

- 1) **Determine the Number of Clusters:** Establishing the value of K or the total number of groups to be formed from the outset.

- 2) Determine the Initial Cluster Centers (Centroids): Randomly selecting initial center points (centroids) for each respective cluster.
- 3) Calculate the Distance to Each Centroid: Measuring the distance of each data object to all cluster center points using a distance formula (e.g., Euclidean Distance).
- 4) Group Objects Based on Minimum Distance: Assigning each data object to the cluster that exhibits the closest/minimum centroid distance to it.
- 5) Do Objects Change Clusters? (Decision/Looping Stage):
- 6) Yes: If any data objects change clusters compared to the previous iteration, the system reverts to step 2 to update the centroids and repeat the calculation process.
- 7) No: If the object positions stabilize and no further transitions occur, the process finishes and proceeds to the next stage.
- 8) Cluster Results: Generating optimally clustered data.

Classification Sub-Process: KNN (Right Box)

- 1) Determine the Number of Neighbors (K): Setting the number of nearest neighbors (K) to serve as a reference for determining the class/group of the test data.
- 2) Calculate Distance from Test Data to Training Data: Computing the distances between the new data (test data) and all existing instances in the training dataset.
- 3) Sort Data Based on the Smallest Distance: Ordering the distance calculation results from the closest (smallest) to the furthest.
- 4) Determine the Test Data Group Based on the Majority Label: Examining the specified K nearest neighbors from the sorted list and assigning the class/group to the test data based on the most frequently occurring group label via majority voting.

3 Result and Discussion

3.1 Result

This section presents the analytical findings of the study conducted by combining the K-Means and KNN methods. The evaluation was carried out by testing the hyperparameter K value of the KNN method on the dataset previously structured by the K-Means method, with the K parameter evaluated within a range from K=3 to 9.

KNN Experimental Results. Figure 2 illustrates the performance evaluation of the classification model using the accuracy metric against variations of the hyperparameter K value (X-axis), ranging from 3 to 9. The Y-axis represents the model's Accuracy Value on a decimal scale (0 to 1.2, where a value of 1 corresponds to 100% accuracy).

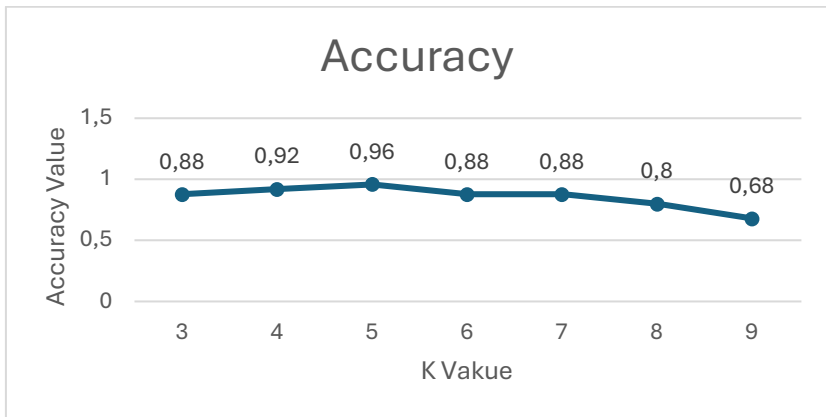


Figure 2. KNN Test Results

The detailed experimental data represented by the graph is outlined in Table 1.

Table 1. KNN Test Results

K Value	Accuracy Value	Accuracy Percentage
3	0.88	88%
4	0.92	92%
5	0.96	96% (Optimal)
6	0.88	88%
7	0.88	88%
8	0.8	80%
9	0.68	68%

Based on the observed graphical trend, several critical analytical points can be inferred

1) Peak Performance (Optimal Accuracy)

The model achieves its maximum (optimal) performance when $K = 5$, yielding an accuracy rate of 0.96 (96%). This indicates that considering 5 nearest neighbors offers the best generalization capability for the model to classify test data accurately.

2) Graphical Trend Characteristics

- Incline Phase ($K = 3$ to $K = 5$):* An upward trajectory in accuracy is observed, climbing from 0.88 to 0.92, and peaking at 0.96. In this phase, increasing the K value assists the model in mitigating the impact of data noise, thereby stabilizing classification decisions.
- Decline Phase ($K = 6$ to $K = 9$):* Beyond the optimal threshold, the model's performance suffers a significant degradation, starting with a stagnant accuracy of 0.88 at $K = 6$ and $K = 7$, followed by a sharp drop to its lowest point of 0.68 (68%) at $K = 9$.

3) Justification of the Underfitting Phenomenon

The sharp decline in accuracy at higher K values (specifically at $K = 8$ and $K = 9$) is theoretically attributed to the underfitting phenomenon (or oversmoothing of the decision boundary). When the K value is excessively large, the model tends to

include neighbors from alternative classes that are spatially distant, thereby obscuring the actual decision boundaries (causing the decision boundary to become overbroad and biased).

From these experimental results, it can be concluded that the configuration of $K = 5$ represents the best parameter to implement in the model, as it yields the highest accuracy while maintaining a robust balance between model stability (preventing overfitting) and classification sharpness (preventing underfitting).

Clustering Results. The clustering evaluation results show that your model has formed a fairly dense and stable cluster structure, where a Davies-Bouldin Index of 0.8413 indicates that the individual clusters are reasonably well-separated, supported by a Calinski-Harabasz Index of 52.6131, which confirms that the inter-cluster variance ratio is validly structured. Meanwhile, a Silhouette Score of 0.4236 indicates that most data objects are placed in the correct group, although there is still a slight area of overlap at the boundaries between clusters. Overall, the performance of this model is considered moderate-to-good and is well-suited to serve as a foundation for analysis, with the primary room for improvement being the separation of data points in the boundary areas.

Based on the pie chart visualization in Figure 3, the research data was successfully grouped into 5 (five) distinct cluster segments: Cluster 0, Cluster 1, Cluster 2, Cluster 3, and Cluster 4. The overall sample size distributed across these five clusters comprises 82 data points. A detailed summary of the proportions and membership counts for the five clusters is provided in Table 2 below.

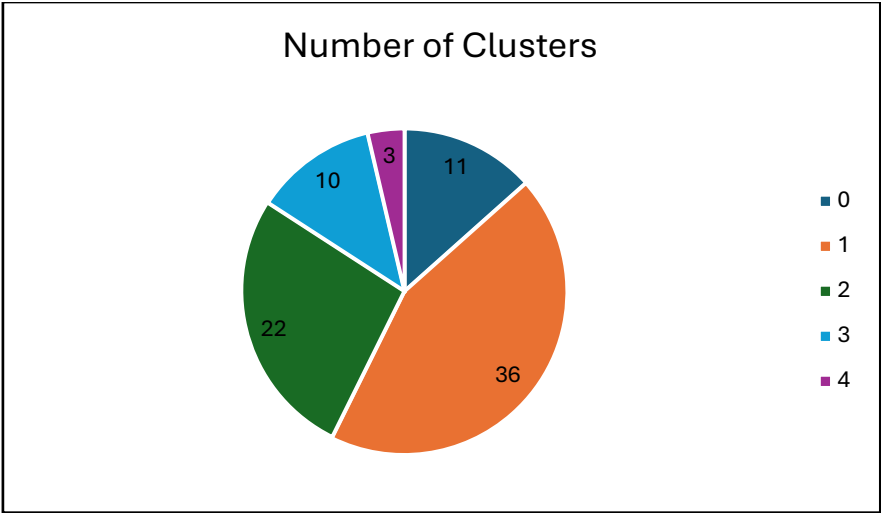


Figure 3. Clustering Results

Table 2. Clustering Results

Cluster Label	Number of Members (Frequency)	Percentage (%)
Cluster 1	36	43,90%
Cluster 2	22	26,83%
Cluster 0	11	13,41%
Cluster 3	10	12,20%
Cluster 4	3	3,66%
Total	82	100,00%

Based on the data clustering results, a detailed summary of the proportions and population sizes for the five clusters is presented in Table 2 below.

1) Majority (Dominant) Cluster

Cluster 1 represents the largest cohort, encompassing 36 members or approximately 43.90% of the total data population. This prominent dominance implies that the characteristics or features inherent to Cluster 1 reflect the general behavior or primary trend of the majority of the investigated objects within this population.

2) Moderate Clusters

Cluster 2 occupies the second-largest position with 22 members (26.83%). Together with Cluster 1, these two cohorts encompass over 70% of the total data, indicating a robust concentration of characteristics within these two segments. Conversely, Cluster 0 (11 members or 13.41%) and Cluster 3 (10 members or 12.20%) exhibit relatively balanced, moderate-to-small group sizes.

3) Minority (Specific/Niche) Cluster

Cluster 4 contains the lowest membership, with only 3 data points (3.66%). In cluster analysis, an exceptionally small group such as this indicates two academic possibilities:

- a) The objects within possess highly unique, specific, and rare characteristics (a niche group) compared to the general population.
- b) This group potentially indicates the presence of outliers that diverge significantly from the traits of the larger cohorts.

KNN Classification Results. The confusion matrix heatmap shows that this 4-class classification model (0, 1, 2, and 3) performs remarkably well, achieving an accuracy of 96% (24 out of 25 samples correctly predicted). The model recorded perfect predictions for Class 1 (9 samples), Class 2 (10 samples), and Class 3 (3 samples) without any errors (100% recall). The only misclassification occurred in Class 0, where out of 3 actual samples, 2 were correctly predicted while the remaining 1 was mistakenly classified as Class 1.

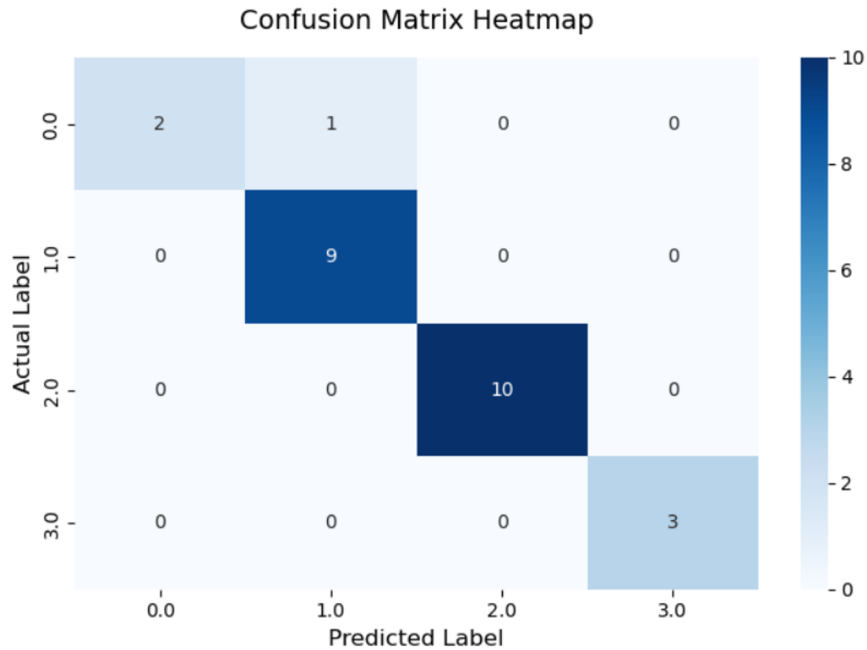


Figure 4. Heatmap Confusion Matrix

Based on the model evaluation using the K-Nearest Neighbors (KNN) method, the classification report presented in Table 3 measures the model's performance in predicting 4 distinct target classes: Class 0.0, Class 1.0, Class 2.0, and Class 3.0. A total of 25 test data samples (support) were utilized in this evaluation process.

Table 3. KNN Classification Report with K=5

	Classification Report			
	Precision	Recall	F1-Score	Support
0.0	1.00	0.67	0.80	3
1.0	0.90	1.00	0.95	9
2.0	1.00	1.00	1.00	10
3.0	1.00	1.00	1.00	3
accuracy			0.96	25

Per-Class Performance Analysis

- a) Class 0.0 (3 Samples)
The model achieves a Precision rate of 1.00, meaning that all instances predicted as Class 0.0 were entirely accurate (100%). However, the Recall value of 0.67 indicates that the model missed approximately 33% of the actual Class 0.0 instances (likely misclassifying them into another class). This imbalance results in an F1-Score of 0.80.
- b) Class 1.0 (9 Samples)

The model successfully identified all actual instances of Class 1.0 without omission, as evidenced by a Recall value of 1.00. Conversely, the Precision rate of 0.90 indicates a slight presence of false positives, where an instance from another class (presumably Class 0.0) was erroneously classified as Class 1.0. Overall, the performance for this class remains remarkably high with an F1-Score of 0.95.

c) Class 2.0 (10 Samples) & Class 3.0 (3 Samples)

Both classes exhibit flawless performance metrics. The Precision, Recall, and F1-Score for these two classes uniformly reach 1.00 (100%). This demonstrates that the KNN algorithm successfully recognized the underlying feature patterns of Class 2.0 and Class 3.0 perfectly, without encountering any misclassifications. In an overarching context, the KNN model demonstrates highly impressive performance on this dataset. The overall Accuracy reaches 0.96 (96%), signifying that out of the 25 provided test instances, the model successfully predicted 24 correctly, committing only a single misclassification. The singular prediction error, appears to be a single sample from Class 0.0 that was mistakenly categorized as Class 1.0. This phenomenon is substantiated by the lower Recall in Class 0.0 (where only 2 out of 3 samples were accurately identified) and the subsequent dip in Class 1.0 Precision to 0.90 due to the accumulation of the single false positive from the other class.

4 Conclusion

Based on the empirical findings and analysis regarding the classification of student academic risks using an Educational Data Mining (EDM) framework driven by a hybrid K-Means + K-Nearest Neighbor (KNN) method, several conclusions can be drawn. The performance evaluations demonstrate that synthesizing these two methods yields a highly impressive accuracy. The hyperparameter configuration of $K = 5$ for KNN proved to be the most optimal architecture, achieving an accuracy of 96%, whereby the model accurately predicted 24 out of 25 test samples.

Expanding the K value beyond 5 acts as a catalyst for the underfitting phenomenon, which blurs the decision boundaries and subsequently degrades the accuracy down to 68% at $K = 9$. In terms of granularity, the model is capable of recognizing the patterns of Class 2.0 and Class 3.0 flawlessly (100%), while registering only a minor misclassification between Class 0.0 and Class 1.0. On the whole, the hybrid K-Means + KNN model has proven to be highly adaptive and accurate. The implementation of this system can serve as a robust foundational framework for higher education institutions to build early warning systems, enabling academic interventions such as counseling or remedial programs to be deployed rapidly and precisely to foster student academic success.

References

- Alfiah, F., Setiadi, A., & Andriyani, F. (2026). Implementasi Algoritma Naive Bayes Pada Sistem Pendukung Keputusan Prediksi Kelulusan Mahasiswa. In *Journal* (Vol. 12, Number 1).
- Ariawan, M. P. A., Ciptayani, P. I., & Peling, I. B. A. (2024). Application of the K-Means Algorithm to Evaluate Basic Programming Abilities of Undergraduate Students in Applied Software Engineering Technology. *Proceedings of the International Conference on Sustainable Green Tourism Applied Science - Engineering Applied Science 2024 (ICoSTAS-EAS 2024)*, 368–377. https://doi.org/10.2991/978-94-6463-587-4_42
- Ariawan, M. P. A., Peling, I. B. A., & Subiksa, G. B. (2023). Prediksi Nilai Akhir Matakuliah Mahasiswa Menggunakan Metode K-Means Clustering (Studi Kasus : Matakuliah Pemrograman Dasar). *Jurnal Nasional Teknologi Dan Sistem Informasi*, 9(2). <https://doi.org/10.25077/teknosi.v9i2.2023.122-131>
- Farouk, A., Zeyad, H. &, & Khalil, A. (2023). Student Performance Prediction Using Machine Learning: A Comprehensive Analysis. *ASRIC Journal on Engineering Sciences*, 4(1), 266276. <https://asric.africa/engineering-sciences>
- Guozhang, L., Alfred, R., Pailus, R., Fengchang, X., & Haviluddin, H. (2025). Scalable and Efficient Student Behavior Prediction using Parallelized Clustering and AHP-weighted KNN. *Journal of ICT Research and Applications*, 19(2), 142–165. <https://doi.org/10.5614/itbj.ict.res.appl.2025.19.2.3>
- Lu, Y., Yeom, S., Maktoubian, J., Rahman, M. M., & Kim, S.-H. (2025). Improve Student Risk Prediction with Clustering Techniques: A Systematic Review in Education Data Mining. *Education Sciences*, 15(12). <https://doi.org/10.3390/educsci15121695>
- Nadia Bakri, S., & Sofinah Harahap, L. (2025). Analisis klasifikasi Algoritma K-Nearest Neighbor (K-NN) pada struktur Daerah di Kota Medan Classification analysis of K-NN (K-Nearest Neighbor) Algorithm on Regional structures in Medan City. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 2025(2), 182–193. <https://jurnal.unity-academy.sch.id/index.php/jirsi/index>
- Patidar, K., Tiwari, D., Saxena, M., Khan, I., Khatarkar, V., & Pandagre, K. (2025). *Performance prediction of students data by using k-means and KNN algorithms*. 20085. <https://doi.org/10.1063/5.0256108>
- Saputra, H., Aldino Uneputty, R., Agustinus Simyapen, L., Farhan Binalawan Mustamir, M., Ayu Ikawanti, F., & Putri Kusumawati, S. (2025). Analisis Pola Kegiatan Belajar Mahasiswa terhadap Keberhasilan Akademik Menggunakan Algoritma Apriori. *Jurnal Sains Informatika Terapan (JSIT) E-ISSN*, 4(2), 143–154.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

