International Journal of Innovative Research and Scientific Studies, 8(7) 2025, pages: 145-153

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Bayesian modeling of student performance dynamics based on LMS interaction data

Bibigul Razakhova[1,3], Yerbol Aldanov[2], Altynbek Sharipbay[1,3], Assel Omarbekova[1,3], Zhanar Lamasheva[3*]

[1]*Public Association "Kazakhstan Academy of Artificial Intelligence, Kazakhstan.*
[2]*Higher School of Information Technology and Engineering, Astana International University, Kabanbay Batyra ave.,8, Astana, 010000, Kazakhstan.*
[3]*Faculty of Information Technologies, L.N. Gumilyov Eurasian National University, Satpayev str. 2, Astana, Kazakhstan.*

Corresponding author: Zhanar Lamasheva (*Email: lamasheva_zhb@enu.kz*)

Abstract

This study proposes a method for assessing college students' academic achievement and predicting their learning outcomes using Bayesian inference techniques. Furthermore, the Bayesian framework provides a new and more flexible approach to statistics. By fitting posterior probability distributions that characterize the uncertainty in information, our analysis can be more robust and more understandable. The experiments use a university's Learning Management System (LMS) data. We analyze student submission patterns and activity levels in detail. Using Markov Chain Monte Carlo (MCMC) simulations, a stochastic model was developed to view changes in learning behavior over time. Synthetic datasets were used to validate the model's predictions. The model also pinpoints critical parameters – such as submission intensity and switch points – influencing academic outcomes. The results demonstrate the effectiveness of Bayesian modeling for forecasting success at various learning stages. We also conclude with some practical recommendations about optimizing the curriculum and providing students with personalized support. This research incorporated the field of learning analytics by using real-world educational data to increase decision-making in higher education and then adding probabilistic methods.

Keywords: Bayesian inference, Learning analytics, Academic performance prediction, LMS data, MCMC simulation, Personalized learning.

DOI: 10.53894/ijirss.v8i7.10419
Funding: This study received no specific financial support.
History: Received: 01 August 2025 / **Revised:** 03 September 2025 / **Accepted:** 05 September 2025 / **Published:** 02 October 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

One of the primary goals of education nowadays is to make the learning process more efficient. Achieving this demands not just an optimal instructional organization but also the introduction of tools to assess them and predict their academic performances objectively. Predictive analytics can accurately predict students' success, which empowers educators to identify and support at-risk learners proactively. This paper presents a classical probabilistic model for academic performance prediction, which uses posterior probability distributions (PPDs) to interpret academic outputs in light of tangible parameters related to students' lives and environments. It is expected that the results will support improvements to the curricular design and contribute to the effectiveness of education and training.

Predicting students' academic performance is a complex yet highly relevant task, particularly given the proliferation of diverse educational formats, including traditional in-person, online (e.g., LMS, MOOC), and blended learning models [1]. Research in this domain focuses on the early identification of students experiencing learning difficulties to enable timely pedagogical interventions and enhance educational effectiveness [2]. Studies have shown that early diagnosis of at-risk students not only reduces dropout rates but also improves the overall quality of education [3].

One active research direction involves the application of explainable artificial intelligence (XAI) tools. For instance, the use of the ExpliClas tool, combined with analysis of Open University Learning Analytics (OULAD) data, has enabled the generation of interpretable forecasts of student academic performance. These insights support data-driven decision-making by instructors and educational administrators [4].

Other studies examine data from online courses and distance learning platforms. For example, a study on teacher training via the MOOC platform employed methods to analyze access frequency to course materials and visualize learning dynamics [5]. Similar techniques are applied in learning management systems (LMS), which not only enhance student engagement but also collect detailed data on learning behaviors - enabling new opportunities for predicting academic performance [6].

Machine learning methods for analyzing student learning activity have gained significant traction. One study utilizing data from the Moodle platform applied a regression model to examine the relationship between academic performance, login frequency, and course completion rates, achieving a prediction accuracy of 60.8% [7]. While adopting a similar approach, this study employs Bayesian inference to assess the impact of written report submission intensity on student performance, offering a probabilistic perspective absent in prior models.

Predictive modeling of academic success links current student behavior to future outcomes and enables early identification of learners requiring additional support. Research shows that early-stage diagnostic measures allow instructors to implement timely corrective interventions, thereby enhancing the effectiveness of instruction [8]. Furthermore, modern educational data analysis techniques can reveal behavioral patterns from historical data, substantially improving prediction accuracy [9].

The advancement of analytical infrastructure in education has enabled the development of early warning systems. For example, one study proposed and validated a hierarchical regression model to assess students' digital proficiency based on school affiliation and gender [10]. Other studies employed factor and regression models to predict academic performance using key parameters such as cognitive abilities, personal characteristics, and cumulative grade point average [11].

Thus, the analysis and prediction of student academic performance remain critical tools for enhancing the educational process. Despite notable progress in machine learning and educational data analytics, improving prediction accuracy remains a pressing challenge. This study introduces a methodology grounded in the classical probabilistic approach, enabling the estimation of successful curriculum completion through posterior probability distributions. The findings may support educational institutions in developing personalized learning strategies and enhancing program effectiveness.

2. The Purpose and Objectives of the Study

This article aims to propose a methodology for assessing student knowledge and predicting learning outcomes using a classical probabilistic model. The core objective is to estimate the probability of academic success within a specific educational program based on the derived posterior probability distribution.

3. Materials and Research Methods

This study utilizes academic performance data of full-time students from the Faculty of Information Technology, obtained from the university's LMS (Learning Management System). The proposed methodology applies a Bayesian approach to assess knowledge and predict learning outcomes. The model estimates the probability of academic success based on posterior probability distributions and can be used to evaluate the effectiveness of educational programs. Insights from the analysis may support program managers in refining curriculum components. Synthetic data generation is performed using the Markov Chain Monte Carlo (MCMC) method, based on the estimated posterior distributions of model parameters [12].

The model can be used to predict future learning outcomes and validate them against historical data. For instance, it can identify which teaching strategies are most effective for specific student groups or determine which instructors perform best in particular subject areas. By simulating educational processes through a stochastic model, institutions can enhance learning outcomes, optimize resource allocation, and improve the overall quality of educational programs.

Unlike traditional frequentist statistical methods, the Bayesian approach treats model parameters as random variables with associated probability distributions, rather than fixed values [13]. Bayesian inference therefore quantifies uncertainty in parameter estimation using probability theory [14]. As a result, Bayesian methods are considered more intuitive and less susceptible to misinterpretation. The key distinctions between Bayesian and frequentist approaches are summarized in Table 1.

Table 1.
Bayesian and frequentist methods.

Frequency method	Bayesian method
Examines the probability of observation data, taking into account the hypothesis	Examines the probability that the hypothesis is correct, taking into account the data
Does not include prior probabilities	Includes prior probabilities
Using p-values (i.e., point estimates), which are the expectation of a long-term frequency	The use of posterior probability distributions (i.e., variability together with joint estimates), which is an expression of the degree of confidence
The parameters of the model are fixed values	The model parameters are random variables
Conclusions depend on the subjectivity of the researcher	Conclusions depend on the subjectivity of the user

In Bayesian methods, probability serves as a measure of uncertainty associated with a given statement. Bayesian inference for academic performance prediction offers a more intuitive alternative to frequentist approaches by integrating prior assumptions with newly observed data to generate posterior distributions of model parameters. With the rise of artificial intelligence (AI), educators increasingly rely on data-driven machine learning (ML) techniques and statistical frameworks to uncover patterns in student performance. In Liu and Chen [15] a decision support system was developed by integrating data mining techniques with Bayesian belief networks to predict student learning outcomes. The system's methodology involves four key steps: applying fuzzy logic to identify influential factors, using data mining to construct an impact diagram, employing machine learning to generate probability tables, and building a predictive model for early course exam scores to guide students in addressing their academic weaknesses. "The authors report a forecasting accuracy of 82.6%, confirming the reliability of the proposed system. In Tempelaar, et al. [11] the effectiveness of integrating assessments of students' practical knowledge with learning management system analytics was demonstrated, offering a deeper understanding of overall course effectiveness. The features derived from this interaction were identified as the most significant predictors of academic success, outperforming traditional indicators such as prior course performance [16-18]. These findings represent a modest yet meaningful advancement in applying Bayesian modeling to the analysis of educational outcomes.

Bayesian inference employs probability to quantify uncertainty in the estimation of model parameters [19]. Unlike frequentist methods, it treats parameters as random variables with associated probability distributions, rather than fixed constants. For example, educational institution (EI) managers cannot fully evaluate learning quality without analyzing all components of the teacher-student interaction-ranging from instruction to assessment – which is often impractical. Even when assessments are tested on large student cohorts, resulting confidence in their quality remains limited. Similarly, Bayesian inference updates beliefs about outcomes as new evidence becomes available, yet full certainty is rarely achieved. Thus, Bayesian probability reflects the degree of confidence in an event's occurrence, offering a natural interpretation of uncertainty in statistical reasoning.

The essence of Bayesian inference can be illustrated through the example of student success rates. Suppose we assume that 60 out of 100 students successfully completed a task at level 1. Bayesian reasoning focuses on refining the degree of confidence in this outcome as new evidence becomes available, enabling more accurate estimation of the likelihood of success as a result of training.

For example, the probability of a large number of messages being submitted via the LMS on a given day may be significantly higher [4, 5, 7, 9, 20]. To build a model, we describe this phenomenon mathematically. It can be assumed that on a certain τ during the observation period, the parameter λ increases sharply. In this case, therefore divide this observation period into two parts: Before τ and part after τ. Such a split is called a switchpoint:

$$P(A|X) = \frac{P(X|A)P(A)}{P(X)} \infty P(X|A)P(A) \quad (1)$$

A posterior probability is the outcome of Bayesian analysis that encapsulates all available knowledge about a given problem, based on the observed data and the chosen model. Within the Bayesian framework, incorporating prior uncertainty acknowledges the possibility that initial assumptions may be incorrect. As new data or evidence is obtained, prior beliefs are updated, making subsequent hypotheses more accurate and less prone to error.

4. Research Results

This study analyzes the dynamics of student activity in a selected undergraduate mathematics course using data from the Learning Management System (LMS). The dataset covers the first half of the second semester and includes records of written assignment submissions from 100 students enrolled in IT-related programs, as illustrated in Figure 1.

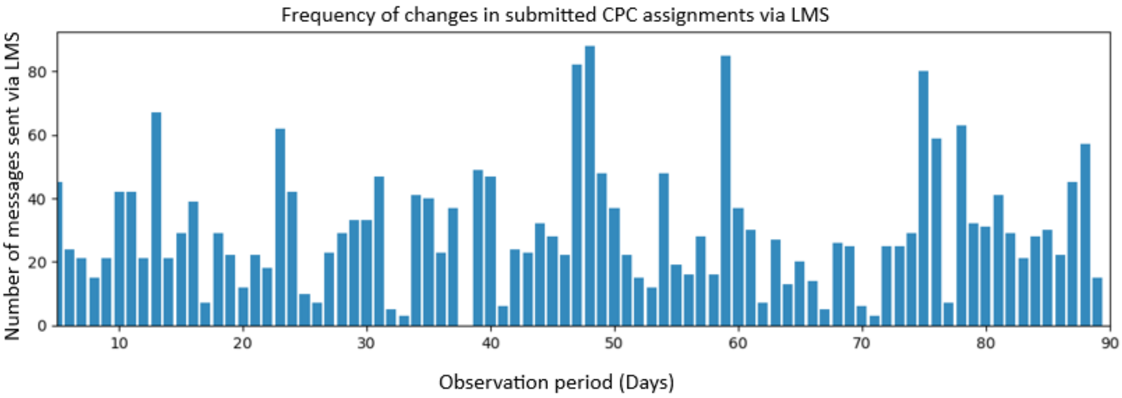

Figure 1.
The intensity of changes in sent written reports.

The objective of this study is to determine whether the nature of the observed data changes over time – and if so, whether the change occurs gradually or abruptly. Typically, task completion and score accumulation reflect a student's engagement with the subject and growing interest in course content. To simulate this scenario, data from 100 students enrolled in the course during the first half of the second academic semester were analyzed. It is assumed that the number of submitted messages represents count data, which can be modeled as a Poisson random variable with a corresponding probability mass function.

$$P(Z=k)=\frac{\lambda^k e^{-\lambda} P(X \mid A)P(A)}{k!};\ k=0,1,2,... \quad (2)$$

It can be assumed denote a random variable M_i – the frequency of the *i*-day with the Poisson law of mass distribution, which is traditionally expressed as:

$$C_i = Poi(\lambda) \quad (3)$$

One of the key challenges is determining the value of the Poisson rate parameter (λ), as its exact value is unknown. As illustrated in Figure 1, certain days – such as days 47, 48, 59, and 76 – exhibit a noticeable increase in the frequency of messages submitted through the LMS. For a Poisson random variable, this is equivalent to an increase in the value of the lambda parameter for these days, the observed periods. It can be assumed that on some day τ during the observation period, the parameter λ increases sharply. In this case, therefore divide this observation period into two parts: before τ and part after τ. Such a split is called a switchpoint:

$$\lambda = \begin{Bmatrix} \lambda_1, & \lambda\, if & t<\tau \\ \lambda_2, & \lambda\, if & t\geq\tau \end{Bmatrix} \quad (4)$$

And if there is no drastic change, then the posterior distribution for the two lambdas will be the same.

For Bayesian inference of unknown lambda parameters, we must select and set a prior probability distributions for various possible values of λ. Since lambda can take any positive values, a continuous density function, such as an exponential distribution function for positive numbers, is suitable as a distribution function of λ_1 and λ_2. It can be assumed write a known density function for an exponential random variable:

$$f_z = (z \mid \lambda) = \lambda \exp(-\lambda z),\ z \geq 0. \quad (5)$$

Therefore, for modeling λ_1 and λ_2, an exponential distribution with a rate parameter α is suitable, which we must certainly include in our model. Also, applying the traditional notation for a random variable with an exponential distribution with the parameter α, we write:

$$\begin{aligned} \lambda_1 &\sim Exp(\alpha) \\ \lambda_2 &\sim Exp(\alpha) \end{aligned} \quad (6)$$

In the case when one parameter depends on another parameter, the latter is called a hyperparameter or a parent variable. In our model, α is a hyperparameter.

The fact that the initial value of the hyperparameter does not greatly affect the model gives us some freedom to choose its value. In Gelman, et al. [21] it is proposed in such a case to set it equal to the inverse of the sample average of the data:

$$\frac{1}{N}\sum_{i=0}^{N} C_i \approx Exp[\lambda \mid \alpha] = \frac{1}{\alpha} \qquad (7)$$

This is explained by the fact that when setting the hyperparameter value according to this formula, we show insignificant confidence in the prior distribution and thereby reduce the influence of the parent variable on its child variable, i.e., on the result.

When the data is noisy, it is a complicated task to choose an prior distribution for the τ(?) branching point. In most cases, when there is a lot of noise in the data, a uniform prior belief of the degree of confidence is well suited, we use this opportunity, and it gives us:

$$\lambda_1 \sim Exp(\alpha), \tau \sim DescreteUniform(1{,}100) \Rightarrow P(\tau = k) = \frac{1}{100}$$

Thus, the final prior distribution for the original random variable will be a superposition of all the distributions obtained above.

Markov Chain Monte-Carlo (MCMC) returns thousands of random variables from the posterior distributions λ_1, λ_2, and τ. Figure 2 shows a histogram based on this data. In Figure 2, we have obtained distributions λ and τ.

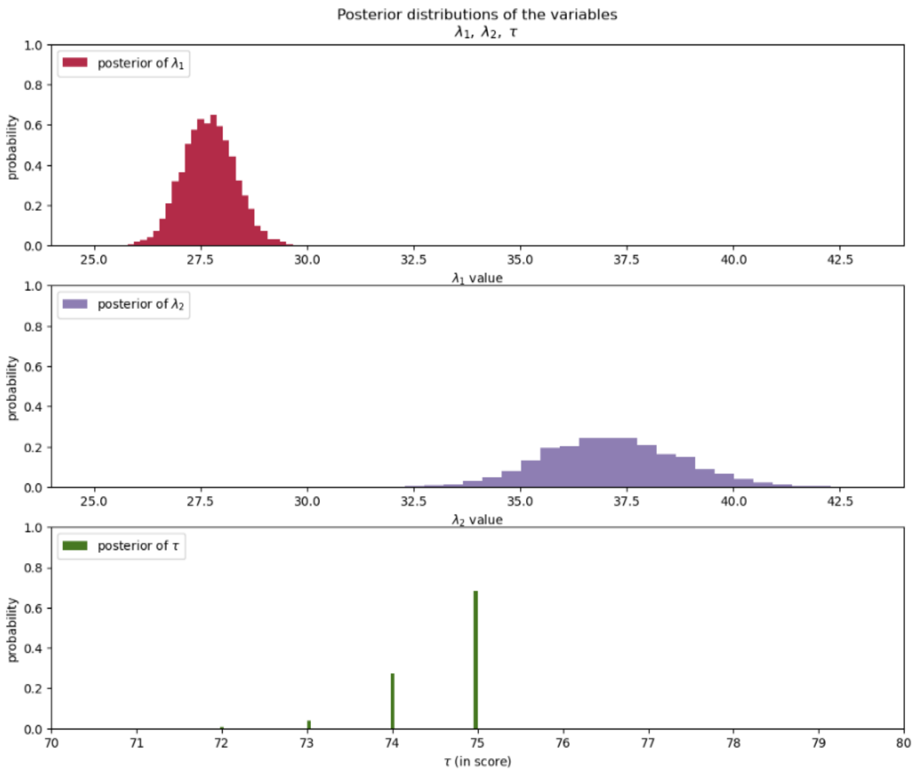

Figure 2.
A posterior parameter distribution.

First, the uncertainty in our estimates is evident: the wider the posterior distribution, the lower the confidence in the parameter values. Subsequently, the most probable values for the parameters can be identified and analyzed.

λ_1 is approximately 27 and λ_2 is approximately 37. The posterior distributions are quite clearly different, indicating a substantially likely change in text messaging. The posterior distribution for tau(?) looks somewhat different than the previous two due to its continuity [22]. Somewhere on the 75th day, the probability of a change in the nature of the intensity of sending the task is 70%. If such a change had not occurred or it turned out to be smoother, then the posterior distribution would have turned out to be smeared because many days would potentially be suitable as a possible one. At the same time, it can be seen from the actual results that only three or four days are somehow suitable for the role of a switchpoint. The resulting forms of posterior distributions, as expected, do not resemble anything from the initial model. This is significantly one of the advantages of the Bayesian approach.

We are also interested in the question of the expected number of completed tasks on day t, $0<t<90$, i.e., the forecast of the expected value of λ at time t?. The average value for all obtained λ_1 and λ_2 are given in Table 2.

Table 2.
Average number of expected messages per day.

Day range	Expected messages
Days 1-72	27.69
Day 73	27.73
Day 74	28.06
Day 75	30.55
Days 76-90	37.13

The results shown in Figure 3 demonstrate the influence of the switch point.

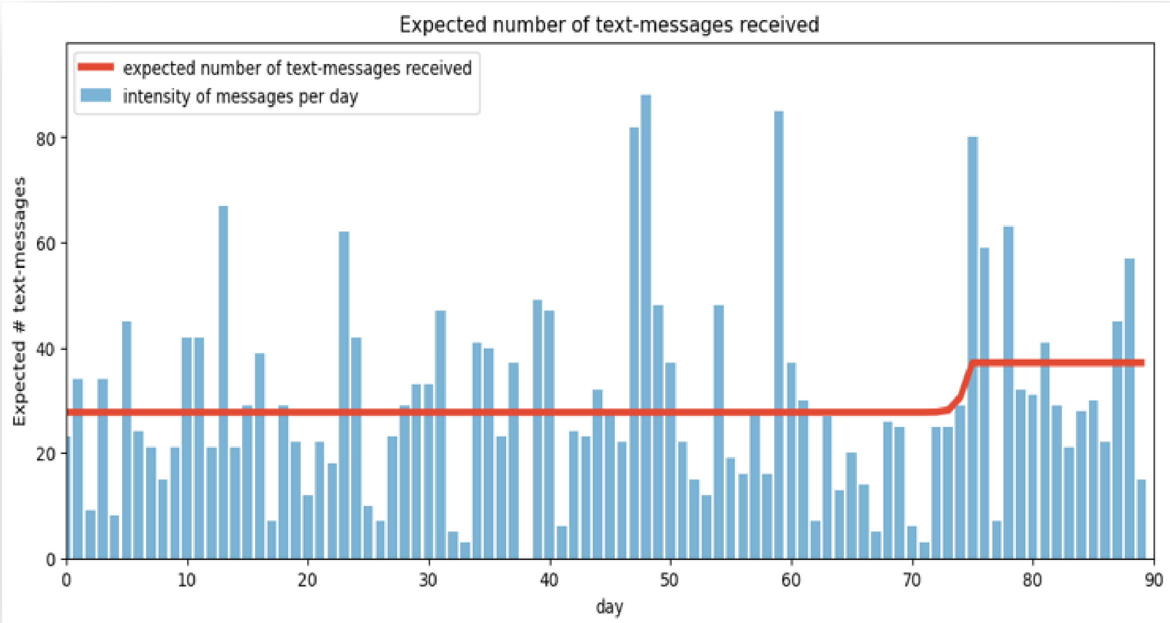

Figure 3.
Comparison of the number of expected shipments with received ones.

In the area of the 75th day, the probability of a change in the intensity of sending the task is 70%. And the results also show that only three or four days are suitable for the role of a branching point

With a complete, even obscure definition of a prior, it is now possible to express the prior distribution of each lambda separately. We will do this for λ_1 (Figure 4).

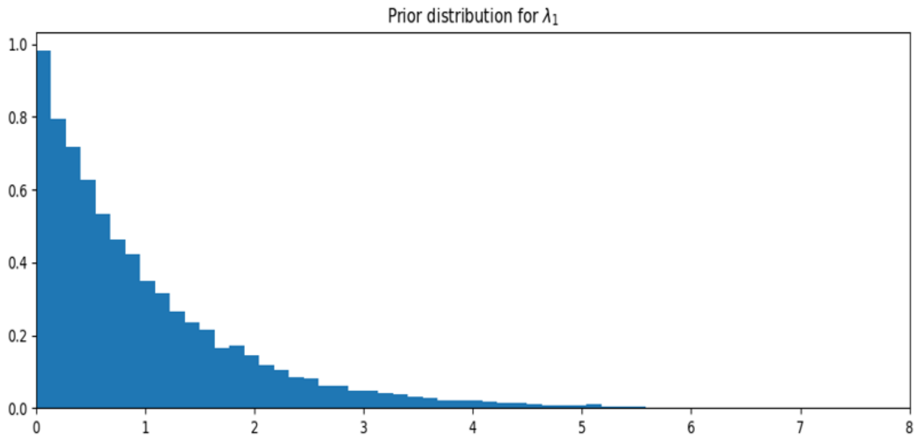

Figure 4.
Example of a prior parameter distribution.

When we set the prior value of ($P(A)$), we have to take into account the data (X) in the model. This can be done by fixing the current value of the variable unchanged, corresponding to the observations that we want to take into account.
Figure 5 shows a visual visualization of the generation of observation data.

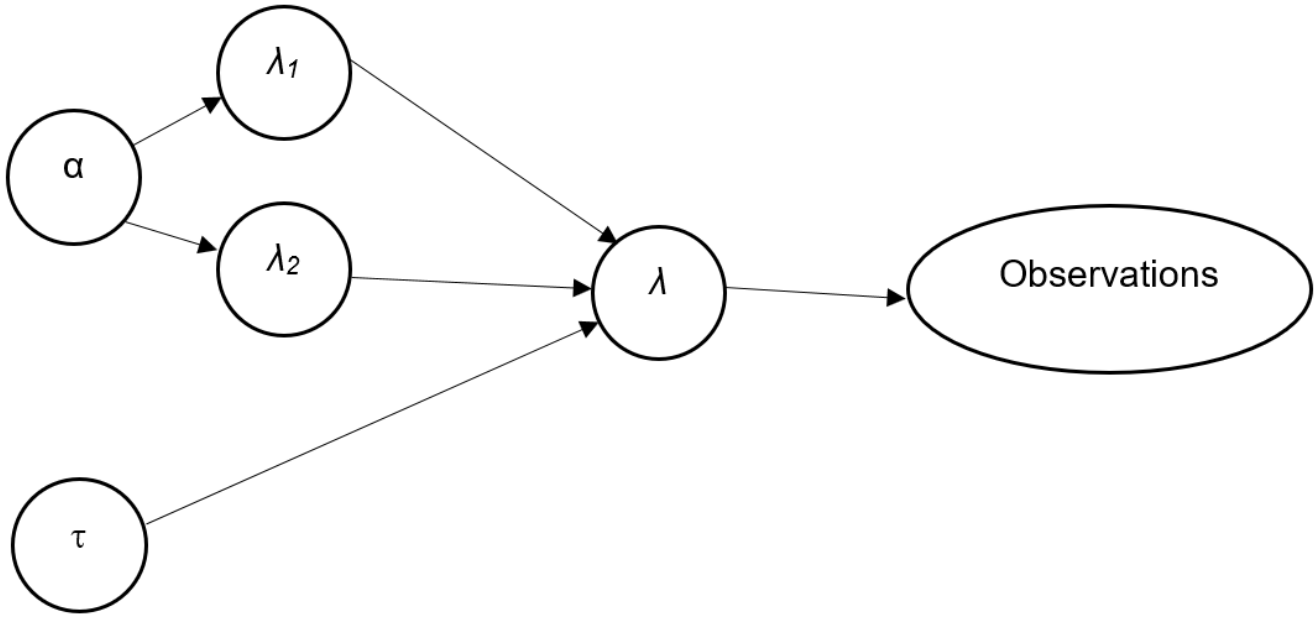

Figure 5.
A visual model of the observation generation scheme.

The Bayesian analysis makes it possible to simulate the probable implementation of our dataset by applying special functions of random non-stochastic variables to generate random results taken from a uniform discrete distribution Khodadadi, et al. [13]. Faes, et al. [23] proposed a multivariate interval approach of inverse uncertainty estimation, enabling safe estimation of model uncertainty when there is sparse experimental data. This possibility of generating artificial data sets is a crucial method of Bayesian inference: making a forecast and checking the suitability of models [14, 24]. The ability to generate artificial data sets is an interesting effect of our modeling. We are creating several more datasets below in Figure 6.

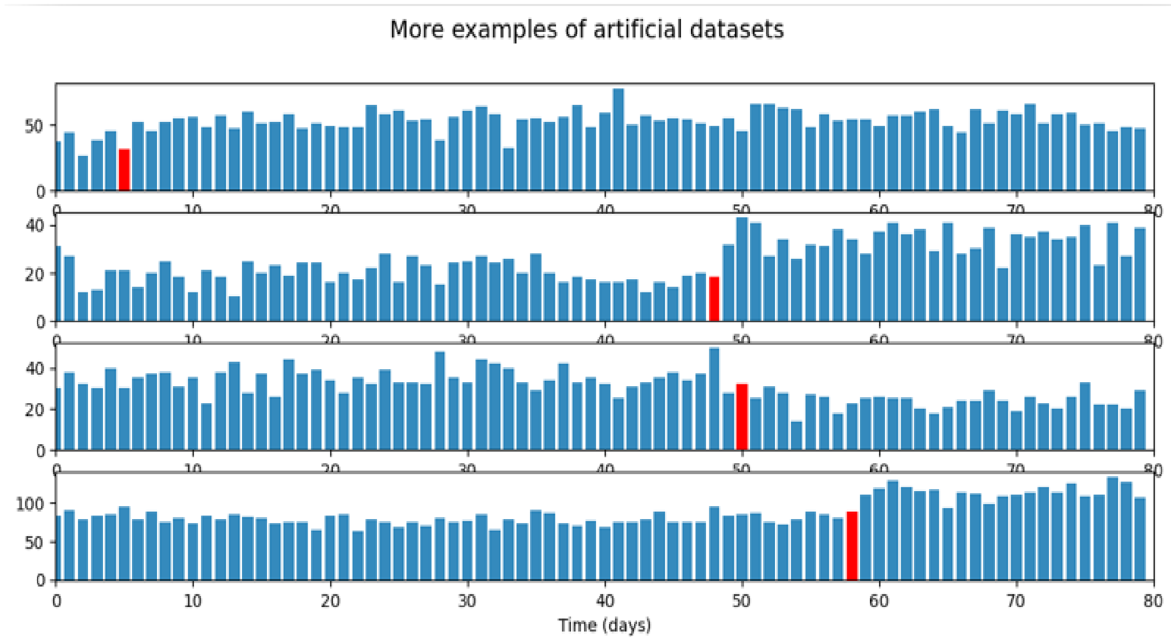

Figure 6.
Generation of artificial data sets.

It's normal that our fictional dataset doesn't look like our observable dataset: the probability that this is the case is incredibly insignificant. A special Markov Chain Monte-Carlo is designed to find good parameters, $\lambda_{1,2}, \tau$, that maximize this probability.

5. Discussions

Some readers who are less familiar with or skeptical of this approach may express concerns about certain aspects of the results. Specifically, the absence of visible uncertainty in the line representing the 'expected number of messages received' may be perceived as a limitation, as some may expect a clearer depiction of variability or confidence intervals. Despite this, our analysis gives serious reasons to it is suggested that student behavior is changing due to a sharp change in the value of λ_1 from λ_2. Otherwise, the value of λ_1 would be closer to λ_2 without a sharp change, which demonstrates a pronounced peak of the posterior distribution. One can only it is hypothesized what led to this change: the complexity of tasks, a student's knowledge decrease, poor quality of educational materials, or new interests in the course of choice.

The authors demonstrated the ability to predict student success based on the intensity of feedback, measured by the number of assignments submitted via the LMS. In the proposed model, submission frequency serves as a key indicator of success at each learning level. A core outcome of the study is the ability to generate predictions using students' historical data -specifically, their prior probability of completing the first level - to construct posterior distributions estimating their likelihood of succeeding at subsequent levels. This approach enables the provision of personalized recommendations tailored to each student's prior achievements.

Furthermore, the integration of semantic technologies and ontology-based models presents a promising direction for enhancing personalized learning. For instance, recent projects have focused on developing smart textbooks in the Kazakh language that generate adaptive learning paths based on the semantic similarity between student responses and structured domain knowledge. Such systems, incorporating fuzzy logic and ontological reasoning, enable automatic assessment and support in real time. These approaches align with the Bayesian framework by allowing uncertainty modeling and individualized learning trajectory construction - extending the applicability of probabilistic models in multilingual and intelligent learning environments.

6. Conclusions

In conclusion, the stochastic model based on the Bayesian approach proves to be an effective tool for evaluating and enhancing educational curricula. It enables the characterization of the probabilistic structure of student progression across various knowledge levels and estimates the likelihood of individual success at each stage of learning. By generating forecasts from posterior distributions using synthetic data, the model assists institutions in student selection and in evaluating academic achievement.

Additionally, it supports the identification of optimal teaching strategies and highlights areas of the curriculum requiring improvement. Ultimately, applying this Bayesian-based stochastic model can lead to improved student outcomes and a significant enhancement in the overall quality of the educational process.

References

[1] S. Kotsiantis, C. Pierrakeas, and P. Pintelas, "Predicting students' performance in distance learning using machine learning techniques," *Applied Artificial Intelligence,* vol. 18, no. 5, pp. 411-426, 2004. https://doi.org/10.1080/08839510490442058

[2] G. W. Dekker, M. Pechenizkiy, and J. M. Vleeshouwers, "Predicting students drop out: A case study," in *Proceedings of the 2nd International Conference on Educational Data Mining, EDM 2009, July 1-3, 2009. Cordoba, Spain*, 2009, pp. 41-50.

[3] M. Baars, L. Wijnia, and F. Paas, "The association between motivation, affect, and self-regulated learning when solving problems," *Frontiers in Psychology,* vol. 8, p. 1346, 2017. https://doi.org/10.3389/fpsyg.2017.01346

[4] J. M. Alonso and G. Casalino, *Explainable artificial intelligence for human-centric data analysis in virtual learning environments. In: Burgos, D., et al. Higher Education Learning Methodologies and Technologies Online. HELMeTO 2019. Communications in Computer and Information Science*. Cham: Springer. https://doi.org/10.1007/978-3-030-31284-8_10, 2019.

[5] S. Jiang, A. Williams, K. Schenke, M. Warschauer, and D. O'Dowd, "Predicting MOOC performance with learning behaviors and NLP-based sentiment analysis," *Educational Data Mining,* vol. 10, no. 1, pp. 46–66, 2014.

[6] H. Drachsler and M. Kalz, "The MOOC and learning analytics innovation cycle (MOLAC): A reflective summary of ongoing research and its challenges," *Journal of Computer Assisted Learning,* vol. 32, no. 3, pp. 281-290, 2016. https://doi.org/10.1111/jcal.12135

[7] A. Jokhan, B. Sharma, and S. Singh, "Early warning system as a predictor for student performance in higher education blended courses," *Studies in Higher Education,* vol. 44, no. 11, pp. 1900-1911, 2019. https://doi.org/10.1080/03075079.2018.1466872

[8] Y. Wang and N. T. Heffernan, *The student skill model. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (Eds.), Intelligent Tutoring Systems. ITS 2012. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-30950-2_51, 2012.

[9] G. Siemens, "Learning analytics: The emergence of a discipline," *American Behavioral Scientist,* vol. 57, no. 10, pp. 1380-1400, 2013. https://doi.org/10.1177/0002764213498851

[10] A. Pardo, R. A. Ellis, and R. A. Calvo, "Combining observational and experiential data to inform the redesign of learning activities " in *Proceedings of the fifth International Conference on Learning Analytics and Knowledge* 2015, pp. 305-309.

[11] D. T. Tempelaar, B. Rienties, and B. Giesbers, "In search for the most informative data for feedback generation: Learning analytics in a data-rich context," *Computers in Human Behavior,* vol. 47, pp. 157-167, 2015. https://doi.org/10.1016/j.chb.2014.05.038

[12] D. Grana, L. De Figueiredo, and K. Mosegaard, "Markov chain Monte Carlo for petrophysical inversion," *Geophysics,* vol. 87, no. 1, pp. M13-M24, 2022. https://doi.org/10.1190/geo2021-0177.1

[13] A. Khodadadi, I. Tsapakis, M. Shirazi, S. Das, and D. Lord, "Derivation of the empirical Bayesian method for the negative Binomial-Lindley generalized linear model with application in traffic safety," *Accident Analysis & Prevention,* vol. 170, p. 106638, 2022. https://doi.org/10.1016/j.aap.2022.106638

[14] A. M. Ellison, "An introduction to Bayesian inference for ecological research and environmental decision-making," *Ecological Applications,* vol. 6, no. 4, pp. 1036-1046, 1996. https://doi.org/10.2307/2269588

[15] K. F. R. Liu and J. S. Chen, "Prediction and assessment of student learning outcomes in calculus a decision support of integrating data mining and Bayesian belief networks," presented at the 2011 3rd International Conference on Computer Research and Development. IEEE. https://doi.org/10.1109/iccrd.2011.5764024, 2011.

[16] S. Iglesias-Pradas, Á. Hernández-García, J. Chaparro-Peláez, and J. L. Prieto, "Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study," *Computers in Human Behavior*, vol. 119, p. 106713, 2021. https://doi.org/10.1016/j.chb.2021.106713

[17] A. Nazyrova, M. Milosz, G. Bekmanova, A. Omarbekova, A. Mukanova, and G. Aimicheva, "Analysis of the consistency of prerequisites and learning outcomes of educational programme courses by using the ontological approach," *Applied Sciences*, vol. 13, no. 4, p. 2661, 2023. https://doi.org/10.3390/app13042661

[18] G. Bekmanova, A. Nazyrova, A. Omarbekova, and A. Sharipbay, *The model of curriculum constructor. In: Gervasi, O., et al. Computational Science and Its Applications – ICCSA 2021. ICCSA 2021. Lecture Notes in Computer Science*. Cham: Springer. https://doi.org/10.1007/978-3-030-86960-1_32, 2021.

[19] R. Van De Schoot *et al.*, "Bayesian statistics and modelling," *Nature Reviews Methods Primers*, vol. 1, p. 1, 2021. https://doi.org/10.1038/s43586-020-00001-2

[20] D. Baneres, M. E. Rodríguez-Gonzalez, and M. Serra, "An early feedback prediction system for learners at-risk within a first-year higher education course," *IEEE Transactions on Learning Technologies*, vol. 12, no. 2, pp. 249-263, 2019. https://doi.org/10.1109/TLT.2019.2912167

[21] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, 3rd ed. Boca Raton, FL: CRC Press, 2013.

[22] H. Waheed, S. U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, "Predicting academic performance at higher education institutions using machine learning models: A systematic literature review," *Computers in Human Behavior*, vol. 104, p. 106179, 2020.

[23] M. Faes *et al.*, "A multivariate interval approach for inverse uncertainty quantification with limited experimental data," *Mechanical Systems and Signal Processing*, vol. 118, pp. 534-548, 2019. https://doi.org/10.1016/j.ymssp.2018.08.050

[24] L. V. Jospin, H. Laga, F. Boussaid, W. Buntine, and M. Bennamoun, "Hands-on Bayesian neural networks—A tutorial for deep learning users," *IEEE Computational Intelligence Magazine*, vol. 17, no. 2, pp. 29-48, 2022.