J. Math. Computer Sci., 34 (2024), 313–325

Online: ISSN 2008-949X

Journal of Mathematics and Computer Science

Journal Homepage: www.isr-publications.com/jmcs

Pre-service mathematics teachers' performance prediction using a modified Mann-Tseng forward-backward splitting algorithm

Nipa Jun-on[a], Pronpat Peeyada[b], Raweerote Suparatulatorn[c,d], Watcharaporn Cholamjiak[b,*]

[a]*Department of Mathematics, Faculty of Science, Lampang Rajabhat University, Lampang 52100, Thailand.*

[b]*School of Science, University of Phayao, Phayao 56000, Thailand.*

[c]*Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand.*

[d]*Office of Research Administration, Chiang Mai University, Chiang Mai 50200, Thailand.*

Abstract

Classifying educational data into a particular category remains challenging due to the massive and extensive number of variables within the dataset. This paper emphasizes a new algorithm for variational inclusion problems with the classification of pre-service mathematics teachers' performance in their method courses through a mathematics teacher education program as its application. First, we propose the modified Mann-Tseng forward-backward splitting algorithm based on inertial technique to speed up the convergence of the algorithm. Then, we prove the weak convergence theorem, we compare and demonstrate the efficacy and applicability of our classification schemes in extreme learning machine (ELM) with other machine learning methods; support vector machine (SVM), logistic regression, boosted trees. Moreover, we compare our algorithm with other algorithms in the same ELM. The application is based on the genuine educational data provided in this paper.

Keywords: Variational inclusion problem, Tseng's method, data classification, method course, pre-service mathematics teacher.

2020 MSC: 47H05, 47J25, 49M37, 97M10, 97C70, 97U70.

1. Introduction

Mathematics teachers who perceive that their pedagogical practices positively impact student learning and achievement are likelier to implement new teaching strategies and take instructional risks [24, 25, 34]. Pre-service teachers enter their teacher education programs with strong beliefs based on experiences as students and interactions with their teachers. In mathematics, where pre-service mathematics teachers frequently have traditional experiences as students, they may enter teacher education programs with a limited understanding of the subject. In order to effectively establish a reform-based environment as teachers, traditionally-minded pre-service mathematics teachers must modify their beliefs regarding mathematics education [9].

*Corresponding author

Email addresses: nipa.676@g.lpru.ac.th (Nipa Jun-on), pronpat.pee@gmail.com (Pronpat Peeyada), raweerote.s@gmail.com (Raweerote Suparatulatorn), watcharaporn.ch@up.ac.th (Watcharaporn Cholamjiak)

doi: 10.22436/jmcs.034.04.01

Received: 2024-01-31 *Revised:* 2024-02-23 *Accepted:* 2024-03-06

However, to educate individuals who will be able to access, examine, and utilize information in the future, it is essential for modern education to provide a high-quality and continuous education. This circumstance necessitates the integration of technologies into instructional environments so that students can manage and construct their own learning processes [16, 21]. Teachers must view technology as an instrument for enhancing the instructional processes of a specific subject, such as mathematics, and not only as a subject of general technology to be taught [21].

Method courses in a teacher education program can help pre-service mathematics teachers acquire the knowledge and abilities necessary to be effective teachers [38]. Method courses in a mathematics teacher education program are designed to provide pre-service teachers with the pedagogical knowledge and skills required to teach mathematics effectively [38]. The emphasis of these courses is on the teaching strategies, methodologies, and best practices for teaching mathematics at different grade levels. By providing a solid foundation in both the content and pedagogy of mathematics, these courses can help pre-service mathematics teachers feel more capable and confident in their roles as teachers. Moreover, method courses can assist them in creating effective lesson plans, employing technology effectively in the classroom, and assessing student progress in a meaningful manner [21, 41, 44].

We hypothesize that the sophistication of teaching mathematics in the modern world necessitates concentrating on pre-service mathematics teachers' performance in method courses. In the digital era, a mathematics teacher's performance in method courses conceivably incorporates mathematical dispositions, self-efficacy for teaching mathematics, and technology-integrated competency.

In this paper, mathematical disposition refers to an individual's mathematical attitudes, beliefs, and affective responses [22]. Pre-service mathematics teachers' mathematical disposition can be influenced and shaped throughout their teacher education program. By promoting positive attitudes, beliefs, and emotional responses towards mathematics, teacher education programs can contribute to improving the academic performance of pre-service mathematics teachers [8].

Furthermore, self-efficacy for teaching mathematics refers to the confidence of pre-service mathematics teachers in their capacity to teach the subject effectively [11, 12]. Self-efficacy for teaching mathematics can be influenced and nurtured through targeted support and experiences provided in teacher education programs. By fostering pre-service teachers' belief in their teaching abilities and providing opportunities to develop and refine their skills, programs can positively impact both the self-efficacy of pre-service teachers and their subsequent academic performance [19].

Also, technology-integrated competency, which refer to the competency to allow students to engage in work on complicated mathematics problems using technology, can have a significant impact on the academic performance of pre-service mathematics teachers in method courses. Technology-integrated competency can empower pre-service mathematics teachers to create dynamic, engaging, and personalized learning experiences [42]. By leveraging technology effectively, they can enhance student engagement, provide individualized support, and make mathematics more accessible and relevant [1].

The educational field has recently been more interested in applying data classification strategies [5]. Classification is a technique for distinguishing the category of provided data elements, also known as targets/labels or categories. It studies discovering new and potentially helpful information or significant outcomes from data. Utilizing various categorization techniques, it also endeavors to discover new trends and patterns in datasets. Specifically, educational data classification is now an effective method for identifying concealed patterns in educational data, predicting students' academic performance, determining teacher competency, and improving the learning and teaching policy plan [5, 17]. In this paper, we focused on pre-service mathematics teachers' information as our educational data for classification to identify latent patterns in their performance in method courses using a modified forward-backward splitting algorithm.

Throughout this article, denote by $\mathcal{H}$ a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. Let $\mathbb{N}$ be the set of nonnegative integers. The problem of identifying a point $x \in \mathcal{H}$ such that

$$0 \in (F + G)x, \tag{1.1}$$

is called the variational inclusion problem (shortly, VIP), where $F : \mathcal{H} \to \mathcal{H}$ is a single valued mapping and $G : \mathcal{H} \to 2^{\mathcal{H}}$ is a multivalued mapping. The VIP (1.1) was known to play a key role in nonlinear analysis involving well-known mathematical models such as composite minimization problems, variational inequality problems, split feasibility problems, and convex programming, with applications in machine learning, signal and image recovery, and more (see [26, 31, 36]). To solve the VIP (1.1), several splitting algorithms have been created and refined; one of the most prominent splitting algorithms is the forward-backward splitting method (see [23, 30] for more information). Chen and Rockafellar [13] used this method in 1997 to obtain a weak convergence result. Later, Tseng [43] created a modification of the forward-backward splitting method, known as the forward-backward-forward method or Tseng's method. This approach makes use of an adaptive line-search rule and relaxes the assumptions of [13] in order to prove weak convergence. Before that in 1964, the inertial extrapolation technique was proposed by Polyak [32] to speed up the convergence of iterative algorithms which is called the heavy ball method. Padcharoen et al. [27] presented a splitting method in 2021 for solving the VIP (1.1) in $\mathcal{H}$, which was developed from Tseng's method with the inertial extrapolation technique. Weak convergence of this method was established under usual assumptions. This method also solved the problems of image deblurring and image recovery.

Tseng [43] introduced the following Tseng's splitting algorithm:

$$\begin{cases} x_0, x_1 \in \mathcal{H}, \\ y_n = J^G_{\mu_n}(I - \mu_n F)x_n, \\ x_{n+1} = P_C(y_n - \mu_n(Fy_n - Fx_n)), \end{cases} \tag{1.2}$$

where μ_n is chosen to be the largest $\mu \in \{\sigma, \sigma l, \sigma l^2, \ldots\}$ satisfying $\mu\|Fy_n - Fx_n\| \leqslant \kappa\|x_n - y_n\|$, $\sigma > 0$, $l \in (0,1)$, and $\kappa \in (0,1)$. It was proved that the sequence $\{x_n\}$ generated by (1.2) converges weakly to an element in $(F+G)^{-1}(0)$.

Padcharoen et al. [27] used the inertial technique with Tseng's splitting algorithm, which they called the modification of Tseng's algorithm. Their algorithm is of the form

$$\begin{cases} x_0, x_1 \in \mathcal{H}, \\ \rho_n = x_n + \delta_n(x_n - x_{n-1}), \\ y_n = J^G_{\mu_n}(I - \mu_n F)\rho_n, \\ x_{n+1} = y_n - \mu_n(Fy_n - F\rho_n), \end{cases} \tag{1.3}$$

where $\{\mu_n\} \subset [a, b] \subset (0, \frac{1}{\ell})$ and $\{\delta_n\} \subset [0, \delta] \subset [0, 1)$. The stepsizes $\{\mu_n\}$ are depending on the Lipschitz constant ℓ.

Recently, Peeyada et al. [31] used the idea of the inertial technique with Mann iteration process and forward-backward algorithm for getting weak convergence for the VIP (1.1) in $\mathcal{H}$, which is called the inertial Mann forward-backward splitting algorithm (IMFBSA):

$$\begin{cases} x_0, x_1 \in \mathcal{H}, \\ \rho_n = x_n + \delta_n(x_n - x_{n-1}), \\ y_n = \rho_n + \eta_n(x_n - \rho_n), \\ x_{n+1} = J^G_{\mu_n}(I - \mu_n F)y_n, \end{cases}$$

where $\{\eta_n \subset (0,1)\}$, $\{\mu_n\} \subset (0, 2\beta)$, and $\{\delta_n\} \subset [0, \infty)$ satisfy the condition such that

$$\text{(C1) } 0 < \liminf_{n\to\infty} \mu_n \leqslant \limsup_{n\to\infty} \mu_n \text{ and (C2) } \sum_{n=1}^{\infty} \delta_n\|x_n - x_{n-1}\| < \infty.$$

This method also solved the signal recovery problem and the data classification problem using the Wisconsin original breast cancer data set as a training set. For some recent results on several algorithms for the VIP (1.1) and related problems, see [14, 18].

In this paper, we present a modified Mann-Tseng forward-backward splitting algorithm for solving problem (1.1) and prove the weak convergence theorem under mild conditions in a real Hilbert space. We apply our main result to solve data classification problem to predict pre-service mathematics teachers' performance. We then compare the performance of our algorithm with Tseng, modified Tseng, and IMFBSA algorithms.

2. Preliminaries

In this section, before we prove our optimization algorithm in Section 3, we give some necessary definitions and lemmas. We denote weak and strong convergence as $\rightharpoonup$ and $\to$, respectively.

Definition 2.1. Let $\mathcal{H}$ be a Hilbert space. A mapping $F : \mathcal{H} \to \mathcal{H}$ is said to be ℓ-Lipschitz continuous if there is $\ell > 0$ such that

$$\|Fx - Fy\| \leqslant \ell\|x - y\|$$

for all $x, y \in \mathcal{H}$.

Definition 2.2. Let $\mathcal{H}$ be a Hilbert space. Let $G : \mathcal{H} \to 2^{\mathcal{H}}$ be a multivalued mapping. Then G is said to be

(i) monotone if for all $(x, u), (y, v) \in graph(G)$ (the graph of mapping G),

$$\langle u - v, x - y\rangle \geqslant 0,$$

(ii) maximal monotone if for every $(x, u) \in \mathcal{H} \times \mathcal{H}$, $\langle u - v, x - y\rangle \geqslant 0$ for all $(y, v) \in graph(G)$ if and only if $(x, u) \in graph(G)$.

Lemma 2.3 ([6])**.** *Let $\mathcal{H}$ be a Hilbert space. Let $F : \mathcal{H} \to \mathcal{H}$ be a mapping and $G : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone mapping. If $T_\mu := (I + \mu G)^{-1}(I - \mu F)$ and $\mu > 0$, then $Fix(T_\mu) = (F + G)^{-1}(0)$, where $Fix(T_\mu)$ is the set of the fixed point of the mapping T_μ.*

Lemma 2.4 ([10])**.** *Let $\mathcal{H}$ be a Hilbert space. If $G : \mathcal{H} \to 2^{\mathcal{H}}$ is a maximal monotone mapping and $F : \mathcal{H} \to \mathcal{H}$ is a Lipschitz continuous and monotone mapping, then the mapping $F + G$ is maximal monotone.*

Lemma 2.5 ([4])**.** *Let $\{a_n\}$ and $\{b_n\}$ be nonnegative sequences of real numbers satisfying $\sum_{n=1}^{\infty} b_n < \infty$ and $a_{n+1} \leqslant a_n + b_n$. Then, $\{a_n\}$ is a convergent sequence.*

Lemma 2.6 ([6, Opial])**.** *Let Ψ be a nonempty set of a Hilbert space $\mathcal{H}$ and $\{x_n\}$ be a sequence in $\mathcal{H}$. Suppose the following assertions hold.*

(i) *For every $x \in \Psi$, the sequence $\{\|x_n - x\|\}$ converges.*
(ii) *Every weak sequential cluster point of $\{x_n\}$ belongs to Ψ.*

Then $\{x_n\}$ converges weakly to a point in Ψ.

3. Main results

To study the convergence analysis, consider the following conditions.

$(\mathcal{C}1)$ $G : \mathcal{H} \to 2^{\mathcal{H}}$ is maximal monotone mapping.
$(\mathcal{C}2)$ $F : \mathcal{H} \to \mathcal{H}$ is ℓ-Lipschitz continuous and monotone mapping.
$(\mathcal{C}3)$ $\Psi := (F + G)^{-1}(0)$ is nonempty.

Algorithm 1

Initialization: Select arbitrary elements $x_0, x_1 \in \mathcal{H}$. Given $\{\eta_n\} \subset [0,1]$. Let $\{\mu_n\} \subset (0, \frac{1}{\ell})$ and $\{\delta_n\} \subset [0,\infty)$ satisfies the condition such that

$$0 < \liminf_{n\to\infty} \mu_n \leqslant \limsup_{n\to\infty} \mu_n < \frac{1}{\ell} \quad \text{and} \quad \sum_{n=1}^{\infty} \delta_n \|x_n - x_{n-1}\| < \infty.$$

Iterative Steps: Construct $\{x_n\}$ by using the following steps:
Step 1. Set

$$\rho_n = x_n + \delta_n(x_n - x_{n-1}) \quad \text{and} \quad z_n = \rho_n + \eta_n(x_n - \rho_n).$$

Step 2. Compute

$$y_n = J^G_{\mu_n}(I - \mu_n F)z_n.$$

If $z_n = y_n$ then stop and $z_n \in \Psi$. Otherwise, go to the next step.
Step 3. Evaluate

$$x_{n+1} = y_n + \mu_n(Fz_n - Fy_n).$$

Replace n by $n+1$ and then repeat **Step 1.**

Remark 3.1. Let $F : \mathcal{H} \to \mathcal{H}$ be a mapping and the item $(\mathcal{C}1)$ holds. According to Lemma 2.3, if $z_n = y_n$ in Algorithm 1, then $z_n \in \Psi$.

We are now ready for the main convergence theorem.

Theorem 3.2. *Let the sequence $\{x_n\}$ generated due to Algorithm 1 and the items $(\mathcal{C}1) - (\mathcal{C}3)$ are satisfied. Then, $\{x_n\}$ converges weakly to a solution of Ψ.*

Proof. Let $\varpi \in \Psi$. From $0 < \liminf_{n\to\infty} \mu_n \leqslant \limsup_{n\to\infty} \mu_n < \frac{1}{\ell}$, there are $n_0 \in \mathbb{N}, \mu > 0$ and $\bar{\mu} < \frac{1}{\ell}$ such that $\mu \leqslant \mu_n \leqslant \bar{\mu}$ for all $n \geqslant n_0$. Next, we prove all following claims.

Claim 1. For any $n \in \mathbb{N}$,

$$\langle y_n - \varpi, y_n - z_n + \mu_n(Fz_n - Fy_n)\rangle \leqslant 0.$$

By using the definition of y_n, we have

$$(I - \mu_n F)z_n \in (I + \mu_n G)y_n.$$

Thus, we can write

$$g_n = \frac{1}{\mu_n}(z_n - y_n - \mu_n Fz_n),$$

where $g_n \in Gy_n$. Since $F + G$ is maximal monotone, we obtain

$$\langle y_n - \varpi, Fy_n + g_n\rangle \geqslant 0,$$

implying that

$$\langle y_n - \varpi, y_n - z_n + \mu_n(Fz_n - Fy_n)\rangle \leqslant 0.$$

Claim 2. For each $n \geqslant n_0$,

$$\|x_{n+1} - \varpi\|^2 + (1 - (\bar{\mu}\ell)^2)\|z_n - y_n\|^2 \leqslant \|z_n - \varpi\|^2.$$

From (1.3), we have

$$\begin{aligned}\|x_{n+1}-\varpi\|^2 &= \|y_n-\varpi+\mu_n(Fz_n-Fy_n)\|^2\\ &= \|(y_n-z_n)+(z_n-\varpi)\|^2+\mu_n^2\|Fz_n-Fy_n\|^2+2\mu_n\langle y_n-\varpi, Fz_n-Fy_n\rangle\\ &= \|y_n-z_n\|^2+\|z_n-\varpi\|^2+\mu_n^2\|Fz_n-Fy_n\|^2+2\langle y_n-z_n, z_n-\varpi\rangle\\ &\quad+2\mu_n\langle y_n-\varpi, Fz_n-Fy_n\rangle\\ &= \|z_n-\varpi\|^2-\|y_n-z_n\|^2+\mu_n^2\|Fz_n-Fy_n\|^2+2\langle y_n-z_n, y_n-\varpi\rangle\\ &\quad+2\mu_n\langle y_n-\varpi, Fz_n-Fy_n\rangle\\ &= \|z_n-\varpi\|^2-\|y_n-z_n\|^2+\mu_n^2\|Fz_n-Fy_n\|^2+2\langle y_n-\varpi, y_n-z_n+\mu_n(Fz_n-Fy_n)\rangle.\end{aligned}$$

Using Claim 1, we get

$$\|x_{n+1}-\varpi\|^2 \leqslant \|z_n-\varpi\|^2-\|y_n-z_n\|^2+\mu_n^2\|Fz_n-Fy_n\|^2.$$

This follows from the Lipschitz continuity of F that, for all $n \geqslant n_0$,

$$\|x_{n+1}-\varpi\|^2 \leqslant \|z_n-\varpi\|^2-\|y_n-z_n\|^2+(\bar{\mu}\ell)^2\|z_n-y_n\|^2.$$

Thus, Claim 2 is established.

Claim 3. $\lim_{n\to\infty}\|x_n-\varpi\| = \lim_{n\to\infty}\|z_n-\varpi\|$.
Indeed, since the definitions of z_n and ρ_n, and using Claim 2, we have, for all $n \geqslant n_0$,

$$\begin{aligned}\|x_{n+1}-\varpi\| &\leqslant \|z_n-\varpi\|\\ &= \|\eta_n(x_n-\varpi)+(1-\eta_n)(\rho_n-\varpi)\|\\ &\leqslant \eta_n\|x_n-\varpi\|+(1-\eta_n)\|x_n-\varpi+\delta_n(x_n-x_{n-1})\|\\ &\leqslant \|x_n-\varpi\|+(1-\eta_n)\delta_n\|x_n-x_{n-1}\|\\ &\leqslant \|x_n-\varpi\|+\delta_n\|x_n-x_{n-1}\|.\end{aligned}$$

Applying this to Lemma 2.5 with $\sum_{n=1}^{\infty}\delta_n\|x_n-x_{n-1}\| < \infty$, we derive the sequence $\{\|x_n-\varpi\|\}$ converges and hence $\lim_{n\to\infty}\|x_n-\varpi\| = \lim_{n\to\infty}\|z_n-\varpi\|$. In particular, $\{x_n\}$ is bounded and also $\{z_n\}$.

Claim 4. $\lim_{n\to\infty}\|x_n-y_n\| = 0$.
By Claim 2, Claim 3 and $\bar{\mu} < \frac{1}{\ell}$, we obtain

$$\lim_{n\to\infty}\|z_n-y_n\| = 0. \tag{3.1}$$

Since the definitions of z_n and ρ_n, and using $\lim_{n\to\infty}\delta_n\|x_n-x_{n-1}\| = 0$, we have

$$\|z_n-x_n\| = (1-\eta_n)\|\rho_n-x_n\| \leqslant \|\rho_n-x_n\| = \delta_n\|x_n-x_{n-1}\| \to 0 \ \text{ as } \ n\to\infty.$$

This together with (3.1) implies that

$$\lim_{n\to\infty}\|x_n-y_n\| = 0.$$

Claim 5. Every weak sequential cluster point of $\{x_n\}$ belongs to Ψ.
Let x^* be a weak sequential cluster point of $\{x_n\}$. Then $x_{n_k} \rightharpoonup x^*$ as $k\to\infty$ for some subsequence $\{x_{n_k}\}$ of $\{x_n\}$. This implies by Claim 4 that $y_{n_k} \rightharpoonup x^*$ as $k\to\infty$. Next, we show that $x^* \in \Psi$. Let $(\nu, u) \in$

$\mathrm{graph}(F+G)$, that is, $u - Fv \in Gv$. It implies by the definition of y_n that $\frac{1}{\mu_{n_k}}(z_{n_k} - y_{n_k} - \mu_{n_k} F z_{n_k}) \in G y_{n_k}$. By the maximal monotonicity of G, we have

$$\left\langle v - y_{n_k}, u - Fv - \frac{1}{\mu_{n_k}}(z_{n_k} - y_{n_k} - \mu_{n_k} F z_{n_k}) \right\rangle \geqslant 0.$$

Thus, by the monotonicity of F, we get

$$\begin{aligned}\langle v - y_{n_k}, u\rangle &\geqslant \left\langle v - y_{n_k}, Fv + \frac{1}{\mu_{n_k}}(z_{n_k} - y_{n_k} - \mu_{n_k} F z_{n_k}) \right\rangle \\ &= \langle v - y_{n_k}, Fv - Fy_{n_k}\rangle + \langle v - y_{n_k}, Fy_{n_k} - Fz_{n_k}\rangle + \frac{1}{\mu_{n_k}}\langle v - y_{n_k}, z_{n_k} - y_{n_k}\rangle \\ &\geqslant \langle v - y_{n_k}, Fy_{n_k} - Fz_{n_k}\rangle + \frac{1}{\mu_{n_k}}\langle v - y_{n_k}, z_{n_k} - y_{n_k}\rangle .\end{aligned}$$

This follows from the Lipschitz continuity of F and (3.1) that

$$\langle v - x^*, u\rangle = \lim_{k\to\infty} \langle v - y_{n_k}, u\rangle \geqslant 0,$$

which, together with the maximal monotonicity of $F+G$, we get that $0 \in (F+G)x^*$. Therefore, $x^* \in \Psi$. Finally, by Lemma 2.6, we can conclude that $\{x_n\}$ converges weakly to a solution of Ψ. □

4. Application to data classification problem

In this section, we focus on extreme learning machine (ELM) [20] for data classification problems to classify pre-service mathematics teachers' academic performance in method courses including seven levels–1-Poor; 2-Below Average; 3-Average; 4-Above Average; 5-Good; 6-Very Good; 7-Excellent. However, none of the pre-service teachers performs at a poor or below-average level. Thus, there are five classess remaining for classification.

ELM for data classification of this problem by defined as follows: The training dataset is defined by $U := \{(x_k, t_k) : x_k \in \mathbb{R}^n, t_k \in \mathbb{R}^m, k = 1, 2, ..., \mathcal{N}\}$ and $\mathcal{N}$ is distinct samples, x_k is an input training data, and t_k is a target. The following output function of single-hidden layer feed forward neural networks (SLFNs) is computed

$$O_k = \sum_{i=1}^{\mathcal{M}} \omega_i \mathcal{A}(g_i x_k + h_i),$$

where $\mathcal{M}$ is hidden nodes, $\mathcal{A}$ is activation function, ω_i is the optimal output weight at the i-th hidden node, g_i is parameter weight at the i-th hidden node, and h_i is bias. The hidden layer output matrix $\mathcal{P}$ is generated as follows:

$$\mathcal{P} = \begin{bmatrix} \mathcal{A}(g_1 x_1 + h_1) & \dots & \mathcal{A}(g_M x_1 + h_M) \\ \vdots & \ddots & \vdots \\ \mathcal{A}(g_1 x_N + h_1) & \dots & \mathcal{A}(g_M x_N + h_M) \end{bmatrix}.$$

To solve ELM is to find optimal output weight $\omega = [\omega_1^T, ..., \omega_M^T]^T$ such that $\mathcal{P}\omega = T$, where $T = [t_1^T, ..., t_N^T]^T$ is the training target data. We can write the solution ω in the from $\omega = \mathcal{P}^\dagger T$, where $\mathcal{P}^\dagger$ is the *Moore-Penrose generalized inverse* of $\mathcal{P}$ does not exist. The regularization of least square problem is considered for good model fitting. This problem can determine as the following convex minimization problem:

$$\min_{\omega\in\mathbb{R}^n} \{\frac{1}{2}\|\mathcal{P}\omega - T\|_2^2 + \sigma\|\omega\|_1\},$$

where σ is a regularization parameter. This problem is called the least absolute shrinkage and selection operator (LASSO) [40].

The classification evaluation, such as accuracy, precision, recall, and F1-score [2] were used to evaluate the performance of the algorithm, shown in (4.1)-(4.4).

$$\text{Precision} = \frac{tp}{tp + fp} \times 100, \tag{4.1}$$

$$\text{Recall} = \frac{tp}{tp + fn} \times 100, \tag{4.2}$$

$$\text{F1-score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100, \tag{4.3}$$

$$\text{Accuracy} = \frac{tp + tn}{tp + fp + tn + fn} \times 100\%, \tag{4.4}$$

where these matrices give True Positive (tp), True Negative (tn), False Positive (fp), and False Negative (fn).

The multi-class cross-entropy loss function is used in multi-class classification. By computing the following average:

$$\text{Loss} = -\sum_{i=1}^{N} y_i \log \hat{y}_i,$$

where $\hat{y}_i$ is the i-th scalar value in the model output, y_i is the corresponding target value, and N is the number of scalar values in the model output.

This data contains 63-instance educational dataset containing 30 attributes relating to mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency. Before starting data training, the overview of the data is shown in the Table 1.

Table 1: The overview of pre-service mathematics teachers' academic performance dataset.

Attribute Name	Mean	Max	Min	SD	CV
1. Positive Attitude	4.3333	5	3	0.6476	14.9441
2. Persistence and Resilience	3.1429	5	1	0.9648	30.6988
3. Curiosity and Inquisitiveness	3.7460	5	2	0.8224	21.9548
4. Confidence and Self-efficacy	2.4762	5	1	1.1196	45.2138
5. Perceived Relevance and Application	4.4286	5	2	0.6890	15.5570
6. Openness to Multiple Approaches	3.9048	5	1	0.9108	23.3245
7. Appreciation of Beauty and Creativity	2.4762	5	1	1.1620	46.9267
8. Growth Mindset	4.5873	5	3	0.5575	12.1507
9. Connections and Application	1.8730	3	1	0.7723	41.2354
10. Reflectiveness and Metacognition	2.6984	4	1	0.8545	31.6665
11. Confidence in Content Knowledge	4.6667	5	3	0.5388	11.5461
12. Pedagogical Confidence	3.0476	5	1	0.9057	29.7180
13. Belief in Teaching Efficacy	2.3810	5	1	1.0988	46.1498
14. Adaptability and Flexibility	3.8253	5	2	0.8336	21.7903
15. Collaboration and Professional Development	2.5397	5	1	1.1046	43.4942
16. Enthusiasm	4.270	5	3	0.7230	16.9336
17. Flexibility and Differentiation	2.4127	5	1	0.9777	40.5245
18. Problem-Solving Orientation	3.8413	5	2	0.8837	23.0043
19. Reflective Practice	2.6667	5	1	0.9672	36.2702
20. Belief in the Value of Mathematics Education	2.9524	5	1	1.2627	42.7697
21. Technological Skills	3.7778	5	1	1.1974	31.6951
22. Integration of Technology	4.0317	5	3	0.7177	17.8014
23. Adapting Instruction	4.1587	5	3	0.7663	18.4274
24. Collaboration and Communication	2.7778	5	1	1.1701	42.1242
25. Digital Citizenship and Safety	4.2698	5	3	0.6275	14.6960
26. Continuous Learning and Growth	4.3492	5	3	0.6515	14.9801
27. Curriculum Alignment	3.8413	5	2	0.9539	24.8323
28. Student-Centered Approach	3.2381	5	1	1.1460	35.3918
29. Adaptability and Troubleshooting Skills	3.4603	5	1	0.9474	27.3794
30. Ethical and Responsible Use	4.2857	5	2	0.8118	18.9415

For finding this result is therefore performed without excluding the most relevant feature from the dataset. Equations (4.1)-(4.4) gives the mathematical representations of the metrics used for results evaluation of Table 2 provide a breakdown of the first results based on the metrics used to evaluate the model. We compared support vector machine (SVM), logistic regression, boosted trees, and our algorithm (ELM) with a smooth data balancing technique selected. The dataset was divided into 80:20 ratios for training and testing by cross-validation with a 5-fold.

Table 2: Comparative analysis of each algorithm.

Algorithm	Precision	Recall	F1-score	Accuracy
SVM	52.63	53.33	52.98	73.00
Logistic Regression	44.4	68.10	53.75	61.90
Boosted Trees	32.35	48.10	38.68	33.30
Our Algorithm (ELM)	53.57	51.67	52.60	78.00

From Table 2, the result shows that the Our Algorithm (ELM) perform better than SVM, Logistic Regression, and Boosted Trees. Besides, the Boosted Trees is the worst classifier for this data.

We set the following parameters for each algorithm:

Table 3: Chosen parameters of each algorithm.

	Tseng	Modified Tseng	IMFBSA	Our Algorithm
μ_n	-	$\frac{1.999}{2(max(eigenvalue(A^T \times A)))}$	$\frac{1.999}{2(max(eigenvalue(A^T \times A)))}$	$\frac{1.999}{2(max(eigenvalue(A^T \times A)))}$
δ_n	-	$\frac{1}{\|x_n - x_{n-1}\|^5 + n^5}$	$\frac{10^{10}}{\|x_n - x_{n-1}\|^3 + n^3 + 10^{10}}$	$\frac{10^{10}}{\|x_n - x_{n-1}\|^5 + n^5 + 10^{10}}$
η_n	-	-	$\frac{1}{2n^2+1}$	$\frac{9n}{100n+1}$
μ	0.7	-	-	-
δ	0.9	-	-	-
τ	0.9	-	-	-

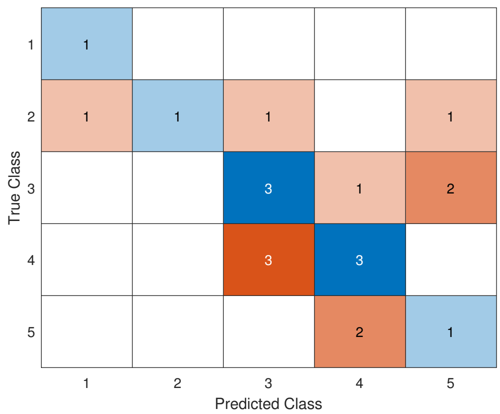

Figure 1: Confusion Matrix.

From Figure 1 shows the confusion matrix of algorithm predictions. True class is on the y-axis and predicted class is on the x-axis. Darker shading indicates higher values.

For comparison, We set sigmoid as an activation function, hidden nodes $\mathcal{M} = 100$ and regularization parameter $\lambda = 0.2$, we obtain the results, as seen in Table 4.

Table 4: The performance of each algorithm.

Algorithm	**Iteration No.**	**Training Time**	**Precision**	**Recall**	**F1-score**	**Accuracy** (%)
Tseng	834	1.1111	54.17	45.00	49.16	76.00
Modified Tseng	81	0.0033	48.57	45.00	46.72	76.00
IMFBSA	33	0.0016	55.00	45.00	49.50	76.00
Our Algorithm	48	0.0063	53.57	51.67	52.60	78.00

Table 4 shows that our algorithm obtained the highest performance accuracy.

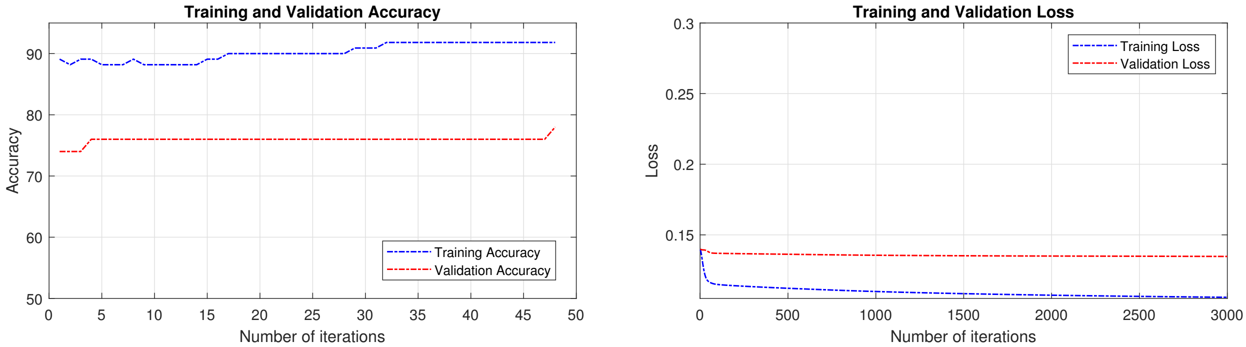

Figure 2: Accuracy and Loss plots of the iteration.

From Figure 2, we can see that the Training Loss and Validation Loss decrease. The Validation Loss value is lower than the Training Loss. On the contrary, when considering the Accuracy graph, it will be found that the Training Accuracy and Validation Accuracy will increase.

We utilized a modified forward-backward splitting algorithm to a 63-instance educational dataset containing 30 attributes relating to mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency—the classification accuracy obtained by our proposed machine learning algorithm. Consequently, 78.00% of the dataset was correctly classified with more accurately than other methods.

5. Conclusion and discussion

This study presents a new method grounded in machine learning algorithms for predicting pre-service mathematics teachers' academic performance levels in method courses, using their data pertaining to various aspects as the source data. The classification accuracy achieved by our proposed machine learning algorithm was applied to a 63-instance educational dataset containing 30 attributes using a modified forward-backward splitting algorithm. As a result, 78.0% of the dataset was classified accurately, which is more accurate to the performance of other methods.

Calculated and compared performances of a modified forward-backward splitting algorithm to predict the academic performance levels, it revealed that attributes relating to mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency acquired throughout the teacher education program were potentially employed to predict pre-service mathematics teachers' academic achievement in method courses.

Mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency can affect pre-service mathematics teachers' academic performance in method courses for several reasons. Firstly, they affect pre-service teachers' motivation and engagement in teaching mathematics. A positive mathematical disposition can enhance pre-service mathematics teachers' motivation and engagement with the subject [15]. When they have a genuine interest and enthusiasm for mathematics, they are more likely to put in the effort to understand the content profoundly and seek out learning opportunities [15]. This motivation and engagement can improve their performance in mathematics method courses [38]. Moreover, pre-service mathematics teachers with high self-efficacy are more likely to exhibit enthusiasm and passion for teaching mathematics [37]. With high self-efficacy of pre-service mathematics teachers, they genuinely believe in the value and importance of mathematics education [35, 39]. When students perceive their teacher's passion and dedication, it can contribute to increased interest and engagement, ultimately impacting their performance when teaching in method courses. Additionally, by embracing technology-integrated competencies, pre-service mathematics teachers can leverage the power of technology to create engaging and compelling learning experiences [29]. Technology allows pre-service mathematics teachers to connect and collaborate with colleagues, educators, and experts worldwide through online communities [21]. These collaborative opportunities provide inspiration, fresh perspectives, and a sense of belonging to a broader community of passionate mathematics teachers.

Secondly, mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency can expand pre-service mathematics teachers' instructional choices and strategies. The mathematical disposition of pre-service mathematics teachers can also influence their instructional choices and strategies used in their method courses [28]. Pre-service mathematics teachers who appreciate the beauty and utility of mathematics are more likely to employ innovative and effective teaching strategies, make connections between diverse topics, and present mathematical concepts meaningfully [3]. Moreover, self-efficacious pre-service mathematics teachers are more likely to believe that they can meet the diverse needs of their students [7, 45]. They have confidence in their ability to adapt their teaching strategies, differentiate instruction, and provide appropriate support to students with varying levels of mathematical proficiency. This adaptability can enhance their effectiveness as teachers and positively impact their academic performance in method courses. Also, technology integration allows pre-service mathematics teachers to leverage various digital tools to enhance their instructional strategies. By effectively utilizing technology, they can create more engaging and interactive lessons catering to different learning styles, positively improving their performance in method courses [33].

Thirdly, pre-service mathematics teachers perform well in mathematics method courses because they are able to foster a positive classroom environment and encourage student engagement. Pre-service mathematics teachers' mathematical disposition can affect the classroom climate they create and their interactions with students [15]. Their positive attitudes and beliefs about mathematics are more likely to nurture a supportive and inclusive learning environment. Furthermore, self-efficacy for teaching mathematics influences the capacity of pre-service teachers to effectively manage their classrooms [37]. A conducive learning environment is created when pre-service mathematics teachers have confidence in their ability to establish a positive and structured learning environment, maintain student engagement, and address discipline issues, effectively. Integrating technology in mathematics instruction can also help pre-service mathematics teachers create learning experiences that are more relevant and captivating for students [33]. Adaptive learning platforms and educational mathematics software can make abstract mathematical concepts more concrete and accessible. When students are actively engaged in the learning process, their motivation and academic performance tend to improve [33]. This situation directly impact to pre-service mathematics teachers' performance in mathematics method courses.

In conclusion, different teacher education programs and institutions may have varying expectations for performance. A good performance in method courses for pre-service mathematics teachers may require a combination of strong content knowledge, effective pedagogical strategies, well-designed lesson plans, assessment skills, and classroom management. However, according to the result of this paper, this combination was affected by attributes relating to mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency. The result can be applied by mathematics teacher educators to predict their students' performance and refine the mathematics teacher education curriculum to focus on mathematical disposition, self-efficacy for teaching mathematics, and technology-integrated competency.

Author's contributions

Conceptualization, Research design, Validation, Analysis, Interpretation, Writing, Review and Editing: N.J.; Writing, Software, Comparison, Data curation: P.P.; Writing, Original Draft, Comparison, Validation, Data curation, Investigation: R.S.; Conceptualization, Research design, Validation, Data curation, Analysis, Application, Review and Editing: W.C.; All authors have read and agreed to the published version of the manuscript.

Acknowledgment

N. Jun-on would like to thank Lampang Rajabhat University for supporting the revenue budget in 2023-2024. W. Cholamjiak and P. Peeyada would like to thank National Research Council of Thailand and University of Phayao (N42A650334), and Thailand Science Research and Innovation, University of Phayao (Fundamental Fund 2024). R. Suparatulatorn would like to thank Chiang Mai Univeristy for partially supporting fund.

References

[1] E. Akman, R. Çakır, *The effect of educational virtual reality game on primary school students' achievement and engagement in mathematics*, Interact. Learn. Environ., **31** (2023), 1467–1484. 1

[2] T. Anwar, S. Zakir, *Deep learning based diagnosis of COVID-19 using chest CT-scan images*, In: 2020 IEEE 23rd International Multitopic Conference (INMIC), IEEE, (2020), 5 pages. 4

[3] K. O. Asante, *Secondary students' attitudes towards mathematics*, Ife Psychol., **20** (2012), 121–133. 5

[4] A. Auslender, M. Teboulle, S. Ben-Tiba, *A logarithmic-quadratic proximal method for variational inequalities*, Comput. Optim. Appl., **12** (1999), 31–40. 2.5

[5] R. S. Baker, K. Yacef, *The state of educational data mining in 2009: A review and future visions*, J. Educ. Data Min., **1** (2009), 3–17. 1

[6] H. H. Bauschke, P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, Springer, New York, (2011). 2.3, 2.6

[7] E. F. Bedel, *Exploring academic motivation, academic self-efficacy and attitudes toward teaching in pre-service early childhood education teachers*, Educ. Stud. Train. Math., **4** (2016), 142–149. 5

[8] J. Beyers, *Development and evaluation of an instrument to assess prospective teachers' dispositions with respect to mathematics*, Int. J. Bus. Soc. Sci., **2** (2011), 20–32. 1

[9] N. Boz, *Turkish pre-service mathematics teachers' beliefs about mathematics teaching*, Aust. J. Teach. Educ., **33** (2008), 66–80. 1

[10] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, (1973). 2.4

[11] A. B. Brown, *Non-traditional preservice teachers and their mathematics efficacy beliefs*, Sch. Sci. Math., **112** (2012), 191–198. 1

[12] C. Y. Charalambous, G. N. Philippou, L. Kyriakides, *Tracing the development of preservice teachers' efficacy beliefs in teaching mathematics during fieldwork*, Educ. Stud. Math., **67** (2008), 125–142. 1

[13] G. H.-G. Chen, R. T. Rockafellar, *Convergence rates in forward–backward splitting*, SIAM J Optim., **7** (1997), 421-444. 1

[14] W. Cholamjiak, P. Cholamjiak, S. Suantai, *An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces*, J. Fixed Point Theory Appl., **20** (2018), 17 pages. 1

[15] J. M. Cruz, A. T. Wilson, X. Wang, *Connections between pre-service teachers' mathematical dispositions and self-efficacy for teaching mathematics*, Int. J. Res. Educ. Sci., **5** (2019), 400–420. 5

[16] L. Darling-Hammond, L. Flook, C. Cook-Harvey, B. Barron, D. Osher, *Implications for educational practice of the science of learning and development*, Appl. Dev. Sci., **24** (2020), 97–140. 1

[17] E. Fernandes, M. Holanda, M. Victorino, V. Borges, R. Carvalho, G. Van Erven, *Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil*, J. Bus. Res., **94** (2019), 335–343. 1

[18] A. Gibali, D. V. Thong, *Tseng type methods for solving inclusion problems and its applications*, Calcolo, **55** (2018), 22 pages. 1

[19] G. Gresham, *Mathematics anxiety and mathematics teacher efficacy in elementary pre-service teachers*, Teach. Educ., **19** (2008), 171–184. 1

[20] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, *Extreme learning machine: theory and applications*, Neurocomputing, **70** (2006), 489–501. 4

[21] N. Jun-on, R. Suparatulatorn, D. Kaewkongpan, C. Suwanreung, *Enhancing pre-service mathematics teachers' technology integrated competence: Cooperative initiation and open lesson observation*, Int. J. Inf. Educ. Tech.,**12** (2022), 1363–1373. 1, 5

[22] I. Kusmaryono, H. Suyitno, D. Dwijanto, N. Dwidayati, *The effect of mathematical disposition on mathematical power formation: Review of dispositional mental functions*, Int. J. Instr., **12** (2019), 343–356. 1

[23] P.-L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., **16** (1979), 964–979. 1

[24] A. Love, A. C. Kruger, *Teacher beliefs and student achievement in urban schools serving African American students*, J. Educ. Res., **99** (2005), 87–98. 1

[25] C. Mischo, K. Maaß, *The effect of teacher beliefs on student competence in mathematical modeling – An intervention study*, J. Educ. Train. Stud., **1** (2013), 19–38. 1

[26] T. Mouktonglang, R. Suparatulatorn, *Inertial hybrid projection methods with selection techniques for split common fixed point problems in Hilbert spaces*, Politehn. Univ. Bucharest Sci. Bull. Ser. A Appl. Math. Phys., **84** (2022), 47–54. 1

[27] A. Padcharoen, D. Kitkuan, W. Kumam, P. Kumam, *Tseng methods with inertial for solving inclusion problems and application to image deblurring and image recovery problems*, Comput. Math. Methods, **3** (2021), 14 pages. 1, 1

[28] M. F. Pajares, *Teachers beliefs and educational research: Cleaning up a messy construct*, Rev. Educ. Res., **62** (1992), 307–332. 5

[29] J. R. Paratore, L. M. O'Brien, L. Jiménez, A. Salinas, C. Ly, *Engaging preservice teachers in integrated study and use of educational media and technology in teaching reading*, Teach. Teach. Educ., **59** (2016), 247–260. 5

[30] G. B. Passty, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert space*, J. Math. Anal. Appl., **72** (1979), 383–390. 1

[31] P. Peeyada, R. Suparatulatorn, W. Cholamjiak, *An inertial Mann forward-backward splitting algorithm of variational inclusion problems and its applications*, Chaos Solitons Fractals, **158** (2022), 7 pages. 1, 1

[32] B. T. Polyak, *Some methods of speeding up the convergence of iteration methods*, USSR Comput. Math. Math. Phys., **4** (1964), 17 pages. 1

[33] R. Powers, W. Blubaugh, *Technology in mathematics education: Preparing teachers for the future*, Contemp. Issues Technol Teach. Educ., **5** (2005), 254–270. 5

[34] J. A. Ross, P. Grey, *School leadership and student achievement: The mediating effects of teacher beliefs*, Can. J. Educ., **29** (2006), 798–822. 1

[35] A. D. Saputro, S. Atun, I. Wilujeng, A. Ariyanto, S. Arifin, *Enhancing pre-service elementary teachers' self-efficacy and critical thinking using problem-based learning*, Eur. J. Educ. Res., **9** (2020), 765–773. 5

[36] Y. Shehu, P. Cholamjiak, *Iterative method with inertial for variational inequalities in Hilbert spaces*, Calcolo, **56** (2019), 21 pages. 1

[37] K. Singh, M. Granville, S. Dika, *Mathematics and science achievement: Effects of motivation, interest, and academic engagement*, J. Educ. Research., **95** (2002), 323–332. 5

[38] M. D. Steele, A. F. Hillen, M. S. Smith, *Developing mathematical knowledge for teaching in a methods course: the case of function*, J. Math. Teach. Educ., **16** (2013), 451–482. 1, 5

[39] T. Temiz, M. S. Topcu, *Preservice teachers' teacher efficacy beliefs and constructivist-based teaching practice*, Eur. J. Psychol. Educ., **28** (2013), 1435–1452. 5

[40] R. Tibshirani, *Regression shrinkage and selection via the lasso*, J. Roy. Statist. Soc. Ser. B, **58** (1996), 267–288. 4

[41] H. S. Tokmak, T. Y. Yelken, G. Y. Konokman, *Pre-service teachers' perceptions on development of their IMD competencies through TPACK-based activities*, J. Educ. Technol. Soc., **16** (2013), 243–256. 1

[42] J. Tondeur, J. Van Braak, G. Sang, J. Voogt, P. Fisser, A. Ottenbreit-Leftwich, *Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence*, Comput. Educ., **59** (2012), 134–144. 1

[43] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM J. Control Optim., **38** (2000), 431–446. 1

[44] X. Yang, F. K. S. Leung, *The relationships among pre- service mathematics teachers' beliefs about mathematics, mathematics teaching, and use of technology in China*, Eurasia J. Math. Sci. Technol. Educ., **11** (2015), 1363–1378. 1

[45] A. L. Zeldin, F. Pajares, *Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers*, Am. Educ. Res. J., **37** (2000), 215–246. 5