RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2024.s001

ScienceAsia 50S (1): 2024: ID 2024s001: **1–7**

A modified projective bi-inertial forward-backward splitting algorithm for detecting bone mineral density

Pentham Yothawut[a], **Pronpat Peeyada**[b], **Wongthawat Liawrungrueang**[c], **Watcharaporn Cholamjiak**[b,*]

[a] Demonstration School, University of Phayao, Phayao 56000 Thailand
[b] Department of Mathematics, School of Science, University of Phayao, Phayao 56000 Thailand
[c] Department of Orthopaedics, School of Medicine, University of Phayao, Phayao 56000 Thailand

*Corresponding author, e-mail: watcharaporn.ch@up.ac.th

Received 14 Feb 2024, Accepted 6 Jul 2024
Available online 10 Nov 2024

ABSTRACT: In this work, we introduce a projective bi-inertial forward-backward splitting algorithm for solving the sum of two monotone operators in a real Hilbert space, one is maximally monotone and the other is Lipschitz continuous. Under standard assumptions, we prove weak convergence theorems of the proposed algorithm. Furthermore, we provide an application for data classification using an extreme learning machine. To gauge the effectiveness of the algorithm, a reliable dataset for bone mineral density prediction was taken from the Harvard Dataverse. Among the algorithms that have been compared, the best performance was obtained with our algorithm in terms of accuracy, precision, recall, and F1-score. The data classification results show that our algorithm is more efficient in handling classification problems.

KEYWORDS: variational inclusion problem, self adaptive, inertial method, weak convergence, data classification

MSC2020: 46C07 90C90

INTRODUCTION

Let $\mathscr{H}$ be real Hilbert spaces with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$ In this paper, we focus on the following variational inclusion problem (VIP). Let $F:\mathscr{H}\to\mathscr{H}$ and $G:\mathscr{H}\to 2^{\mathscr{H}}$ be, respectively, single-valued monotone mapping and set-valued monotone mapping, the VIP is to:

$$\text{find } x\in\mathscr{H} \text{ such that } 0\in(Fx+Gx), \tag{1}$$

The VIP (1) has significant practical implications across various real-world domains, including engineering, transportation, economics, signal recovery, image processing, and machine learning [1–3]. This broad applicability is substantiated by the research contributions of notable scholars, see in [4–7]. Notably, there is a growing scholarly endeavour dedicated to exploring innovative methodologies for effectively tackling the complexities inherent in the VIP. This research aims to develop robust solutions for this multifaceted problem. The forward-backward splitting method is a famous one of algorithms solving the VIP (refer to [8–10]), which is defined as follows:

$$x^{k+1}=J_{\lambda}^{G}(I-\lambda F)x^{k},\quad k\geqslant 1,$$

where $J_{\lambda}^{G}=(I+\lambda G)^{-1}$ with $\lambda>0$. Furthermore, researchers have enhanced these methods not only to increase their versatility through the use of relaxation techniques (see [11, 12]) but also to enhance their acceleration using inertial techniques (see [13–16]). Later, Alvarez and Attouch [15] further developed the inertial concept to speed up the convergence of an algorithm. The inertial forward-backward method (IFBM) was introduced which defined as follows:

$$\begin{cases} y^{k}=x^{k}+\theta^{k}(x^{k}-x^{k-1}) \\ x^{k+1}=J_{\lambda^{k}}^{G}(I-\lambda^{k}F)y^{k}, \quad k\geqslant 1. \end{cases}$$

The technique for accelerating this method involves the term $\theta^{k}(x^{k}-x^{k-1})$. The forward-backward method with the inertial term θ^{k} for the VIP is a well-established practice, and more comprehensive information can be found in [17–19]. Additionally, in the realm of monotone inclusions and non-smooth convex minimization problems, a convergence theorem has been formally established.

In 2022, Iyiola and Shehu [20] introduced and studied the following two-point inertial proximal point algorithm (TPIPA) for monotone operators in Hilbert spaces:

$$\begin{cases} y^{k}=x^{k}+\theta^{k}(x^{k}-x^{k-1})+\delta^{k}(x^{k-1}-x^{k-2}), \\ x^{k+1}=(1-\alpha^{k})y^{k}+\alpha^{k}J_{\lambda^{k}}^{G}(y^{k}), \end{cases}$$

where $\lambda^{k}>0$. With the suitable conditions of the parameters θ^{k} and δ^{k}, weak convergence theorem was established. The advantage gained with the introduction of $\delta^{k}\in(-\infty,0]$ was shown in the numerical experiments.

Inspired and driven by the research above, we introduce a projective bi-inertial forward-backward splitting algorithm to solve the sum of two monotone operators in a real Hilbert space. We establish a weak convergence theorem for the sequence generated

by our algorithm, subject to appropriate conditions. Additionally, we apply our main result to address a data classification problem related to predicting bone mineral density. To ascertain which classification model offers a more precise dataset prediction, we compute and compare several performance metrics, including accuracy, precision, recall, and F1-score. The results conclusively demonstrate that the proposed algorithm exhibits superior efficiency in handling classification problems.

PRELIMINARIES

In this section, let $\mathcal{H}$ be a real Hilbert space. The weak convergence and strong convergence of $\{x^k\}_{k=1}^{\infty}$ to x are represented by $x^k \rightharpoonup x$ and $x^k \to x$, respectively. We will discuss several crucial fundamental concepts and lemmas in the main results section.

Definition 1 A mapping $F : \mathcal{H} \to \mathcal{H}$ is said to be
(i) monotone mapping if the following holds:

$$\langle Fx - Fy, x - y\rangle \geqslant 0;$$

(ii) L-Lipschitz continuous if there is a constant $L > 0$ such that:

$$\|Fx - Fy\| \leqslant L\|x - y\|;$$

if $L = 1$, then F is called nonexpansive;
(iii) firmly nonexpansive if

$$\|Fx - Fy\|^2 \leqslant \|x - y\|^2 - \|(I - F)x - (I - F)y\|^2,$$

or equivalently

$$\langle Fx - Fy, x - y\rangle \geqslant \|Fx - Fy\|^2;$$

(iv) τ-cocoercive or τ-inverse strongly monotone if τF is firmly nonexpansive when $\tau > 0$.

Lemma 1 ([21]) *Let $F : \mathcal{H} \to \mathcal{H}$ be a nonexpansive mapping such that $Fix(F) \neq \phi$. If there exists a sequence $\{x^k\}$ in $\mathcal{H}$ such that $x^k \rightharpoonup x \in \mathcal{H}$ and $\|x^k - Fx^k\| \to 0$, then $x \in Fix(F)$. Here, $Fix(F) = \{x \in \mathcal{H} : x = F(x)\}$.*

Lemma 2 ([22]) *Let $F : \mathcal{H} \to \mathcal{H}$ be τ-cocoercive mapping and $G : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone mapping. Then, we have*
(i) *for $\lambda > 0$, $Fix(J_{\lambda}^{G}(I - \lambda F)) = (F + G)^{-1}(0)$;*
(ii) *for $0 < \lambda < \overline{\lambda}$ and $x \in \mathcal{H}$,*

$$\|x - J_{\lambda}^{G}(I - \lambda F)x\| \leqslant 2\|x - J_{\overline{\lambda}}^{G}(I - \overline{\lambda}F)x\|.$$

Lemma 3 ([23]) *Let Ω be a nonempty subset of $\mathcal{H}$ and $\{x^k\}$ be a sequence in $\mathcal{H}$. Assume that the following conditions hold.*
(i) *For every $x \in \Omega$, the sequence $\{\|x^k - x\|\}$ converges.*
(ii) *Every weak sequential cluster point of $\{x^k\}$ belongs to Ω.*

Then $\{x^k\}$ weakly converges to a point in Ω.

Lemma 4 ([24]) *Suppose that $\{\mu^k\}$, $\{\delta^k\}$ and $\{\alpha^k\}$ are sequences in $[0, \infty)$ such that $\mu^{k+1} \leqslant \mu^k + \alpha^k(\mu^k - \mu^{k-1}) + \delta^k$, $\forall k \geqslant 1$, $\sum_{k=1}^{\infty} \delta^k < \infty$, and there is $\alpha \in \mathbb{R}$ with $0 \leqslant \alpha^k < \alpha < 1$, $\forall k \geqslant 1$. Then the following conditions are satisfied:*
(i) *$\sum[\mu^k - \mu^{k-1}]_+ < \infty$, where $[s]_+ = \max\{s, 0\}$;*
(ii) *there exists $\mu^* \in [0, \infty)$ such that $\lim_{k\to\infty} \mu^k = \mu^*$.*

MAIN RESULTS

Throughout the paper, we suppose that that E is a nonempty closed and convex subset of $\mathcal{H}$. Let $F : \mathcal{H} \to \mathcal{H}$ be a τ - inverse strongly monotone mapping and $G : \mathcal{H} \to 2^{\mathcal{H}}$ be a set-valued maximal monotone mapping such that $(F + G)^{-1}(0) \cap E \neq \emptyset$.

Algorithm 1 The projective bi-inertial forward-backward splitting algorithm for the VIP

Initialization: Set $\{\alpha^k\} \subset (0, 1)$, $\{\lambda^k\} \subset (0, 2\tau)$, $\{\theta^k\}, \{\delta^k\} \subset (-\infty, \infty)$ and let $x^{-1}, x^0, x^1 \in \mathcal{H}$.

Iterative Steps: Construct $\{x^k\}$ by using the following steps:
Step 1. Compute

$$y^k = (1 - \alpha^k)x^k + \alpha^k J^k x^k$$

Step 2. Compute

$$x^{k+1} = P_E(y^k + \theta^k(x^k - x^{k-1}) + \delta^k(x^{k-1} - x^{k-2})),$$

where $J^k = J_{\lambda^k}^{G}(I - \lambda^k F)$. Set $k = k + 1$ and return to Step 1.

Assumption 1 (i) $\sum_{k=1}^{\infty} |\theta^k| \|x^k - x^{k-1}\| < \infty$ *and* $\sum_{k=1}^{\infty} |\delta^k| \|x^{k-1} - x^{k-2}\| < \infty$;
(ii) $\liminf_{k\to\infty} \alpha^k > 0$;
(iii) $0 < \liminf_{k\to\infty} \lambda^k \leqslant \limsup_{k\to\infty} \lambda^k < 2\tau$.

Theorem 1 *Let $\{x^k\}$ be generated by Algorithm 1 when Assumption 1 hold. Then $\{x^k\}$ converges weakly to an element of $(F + G)^{-1}(0) \cap E$.*

Proof: Let $x^* \in (F + G)^{-1}(0) \cap E$. For each $k \in N$, since J^k is nonexpansive when $\{x^k\} \subset (0, 2\tau)$, we have

$$\begin{aligned}
&\|x^{k+1} - x^*\| \\
&\quad = \|P_E(y^k + \theta^k(x^k - x^{k-1}) + \delta^k(x^{k-1} - x^{k-2})) - x^*\| \\
&\quad \leqslant \|y^k - x^*\| + \theta^k\|x^k - x^{k-1}\| + \delta^k\|x^{k-1} - x^{k-2}\| \\
&\quad \leqslant (1 - \alpha^k)\|x^k - x^*\| + \alpha^k\|J^k x^k - x^*\| \\
&\qquad + \theta^k\|x^k - x^{k-1}\| + \delta^k\|x^{k-1} - x^{k-2}\| \\
&\quad \leqslant (1 - \alpha^k)\|x^k - x^*\| + \alpha^k\|x^k - x^*\| \\
&\qquad + \theta^k\|x^k - x^{k-1}\| + \delta^k\|x^{k-1} - x^{k-2}\| \\
&\quad = \|x^k - x^*\| + \theta^k\|x^k - x^{k-1}\| + \delta^k\|x^{k-1} - x^{k-2}\|.
\end{aligned}$$

By the conditions of (i), it follows from Lemma 4 that $\lim_{k\to\infty}\|x^k - x^*\|$ exists. This implies that $\{x^k\}$ is bounded. Since $J^G_{\lambda^k}$ is a firmly nonexpansive mapping, let $M^k = \theta^k(x^k - x^{k-1}) + \delta^k(x^{k-1} - x^{k-2})$ we have

$$\begin{aligned}
\|x^{k+1}-x^*\|^2 &= \|P_E(y^k+M^k)-x^*\|^2\\
&\leqslant \|y^k-x^*+M^k\|^2\\
&\leqslant \|y^k-x^*\|^2+2\langle M^k, y^k+M^k-x^*\rangle\\
&\leqslant (1-\alpha^k)\|x^k-x^*\|^2+\alpha^k\|J^kx^k-x^*\|^2+2\langle M^k, y^k+M^k-x^*\rangle\\
&\leqslant (1-\alpha^k)\|x^k-x^*\|^2+\alpha^k\big[\|x^k-\lambda^kFx^k-x^*+\lambda^kFx^k\|^2\\
&\quad -\|x^k-\lambda^kFx^k-J^kx^k-x^*+\lambda^kFx^*+J^kx^*\|^2\big]\\
&\quad +2\langle M^k, y^k+M^k-x^*\rangle\\
&\leqslant \|x^k-x^*\|^2+\alpha^k(\lambda^k)^2\|Fx^k-Fx^*\|^2-2\tau\alpha^k\lambda^k\|Fx^k-Fx^*\|^2\\
&\quad -\alpha^k\|x^k-\lambda^k(Fx^k-Fx^*-J^kx^k\|^2+2\langle M^k, y^k+M^k-x^*\rangle\\
&= \|x^k-x^*\|^2-\alpha^k\lambda^k(2\tau-\lambda^k)\|Fx^k-Fx^*\|^2\\
&\quad -\alpha^k\|x^k-\lambda^k(Fx^k-Fx^*-J^kx^k\|^2+2\langle M^k, y^k+M^k-x^*\rangle. \qquad (2)
\end{aligned}$$

Again by the conditions (i)–(iii) and (2), we have

$$\lim_{k\to\infty}\|Fx^k-Fx^*\| = \lim_{k\to\infty}\|x^k-J^kx^k-\lambda^k(Fx^k-Fx^*)\| = 0.$$

This implies that

$$\lim_{k\to\infty}\|x^k-J^kx^k\| = 0. \qquad (3)$$

Since $\liminf_{k\to\infty}\lambda^k > 0$, there is $\lambda > 0$ such that $\lambda^k > \lambda$. By Lemma 2(ii), we obtain

$$\lim_{k\to\infty}\|x^k-J^G_\lambda(I-\lambda F)x^k\| \leqslant \lim_{k\to\infty}\|x^k-J^kx^k\| = 0. \qquad (4)$$

Since $\{x^k\}$ is bounded, we have a weak sequential cluster point $\bar{x}$ of $\{x^k\}$. Using Lemma 1, we get that $\bar{x} \in Fix(J^G_\lambda(I-\lambda F)) = (F+G)^{-1}(0)$. Since E is closed, $\bar{x} \in E$. This implies that $\bar{x} \in (F+G)^{-1}(0)\cap E$. By Lemma 3, we can obtain that $\{x^k\}$ converges weakly to an element $(F+G)^{-1}(0)\cap E$. □

APPLICATION

Osteoporosis is a global public health problem characterized by a metabolic bone-related disorder involving low bone quantity and altered bone microarchitecture, leading to increased bone fragility and a higher risk of fractures. The prevention and screening of the disease are important. Recent research reported that osteoporosis could be prevented by eliminating accumulation of reactive oxygen species which damage osteoblast formation [25]. Early screening in patients is essential as it enables timely medical treatment to prevent undesirable outcomes. The gold standard for diagnosing osteoporosis is detecting bone mineral density (BMD) using dual X-ray absorptiometry; BMD measurement is exact and can predict osteoporotic fractures. However, this approach faces challenges due to a scarcity of doctors with specialized expertise in analyzing such disease conditions, coupled with the complex interaction of risk factors, making diagnosis difficult. Therefore, we demonstrate the effectiveness of our algorithm in detecting BMD. We use the bone mineral density from the publicly Harvard Dataverse, available on the online website [26], to evaluate the proposed algorithm. This dataset contains 38 attributes that indicate characteristics across various age groups over 1,463 records, from young patients to the elderly. According to the World Health Organization (WHO) guidelines, osteoporosis severity is commonly classified into stages based on bone density and fracture risk, typically using a measurement called T-scores. The four levels of osteoporosis severity are: (i) Normal: T-score is above −1.0, indicating normal bone density; (ii) Osteopenia: T-score is between −1.0 and −2.5, indicating lower bone density than normal but not yet classified as osteoporosis; (iii) Osteoporosis: T-score is −2.5 or lower, signifying significantly reduced bone density and an increased risk of fractures; (iv) Severe Osteoporosis: in some classifications, there may be a further stage called “severe osteoporosis”,

Table 1 Overview of osteoporosis dataset.

Attribute name	Max	Min	Mean	Median	SD
Gender	2	1	1.3876	1	0.4874
Age	99.8	28.6	59.8723	57	12.9451
Height	186	141	165.7707	167	8.0737
Weight	113	23	67.1015	67	12.0150
BMI	37.26	9.21	24.3087	24.22	3.3249
L1.4T	6	−3.4	−0.5467	−0.7	1.5185
FNT	2.7	−5.05	−1.2957	−1.35	1.1172
TLT	3.1	−4.8	−0.9214	−1	1.1548
ALT	181	4	23.3849	19	16.5490
AST	128	9	22.6097	21	9.3731
BUN	69.8	1.74	5.5982	21	9.3731
CREA	381.2	5.86	73.9036	70.6	25.6525
URIC	745.3	5.46	348.3645	340.1	96.7824
FBG	24.65	3.13	5.3327	4.96	1.5545
HDL-C	5.46	0.45	1.2520	1.19	0.3806
LDL-C	6.65	0.14	2.5936	2.55	0.8995
Ca	5.84	1.78	2.2371	2.23	0.1618
P	4.41	0.56	1.0347	1.02	0.2035
Mg	1.73	0.097	0.8686	0.87	0.0956
Calsium	1	0	0.1476	0	0.3549
Calcitriol	1	0	0.1723	0	0.3777
Bisphosph	1	0	0.0608	0	0.2391
Calcitonin	1	0	0.0589	0	0.2353
HTN	1	0	0.5530	1	0.4974
COPD	1	0	0.2502	0	0.4333
DM	1	0	0.3322	0	0.4712
Hyperlipid	1	0	0.3917	0	0.4883
Hyperuricemia	1	0	0.1736	0	0.3789
AS	1	0	0.7533	1	0.4313
VT	1	0	0.0191	0	0.1371
VD	1	0	0.0752	0	0.2638
OP	1	0	0.3691	0	0.4827
CAD	1	0	0.2057	0	0.4044
CKD	1	0	0.0383	0	0.1919
Smoking	1	0	0.2556	0	0.0436
Drinking	1	0	0.2269	0	0.4190
Fracture	1	0	0.0185	0	0.1346

Table 2 Setting operators of the algorithms to solve all of the convex minimization problems (5)–(8).

Problem	Setting operators of the algorithms
RLSPL_1	$F(\beta) \equiv \nabla(\frac{1}{2}\|H\beta - T\|_2^2)$, $G(\beta) \equiv \partial(\lambda\|\beta\|_1)$, $E = \mathscr{H}$
RLSPL_2	$F(\beta) \equiv \nabla(\frac{1}{2}\|H\beta - T\|_2^2)$, $G(\beta) \equiv \partial(\lambda\|\beta\|_2^2)$, $E = \mathscr{H}$
RLSPCL_1	$F(\beta) \equiv \nabla(\frac{1}{2}\|H\beta - T\|_2^2)$, $G(\beta) \equiv \partial(\lambda\|\beta\|_1)$, $E = \{\beta \in \mathbb{R}^M : \|\beta\|_1 \leqslant \rho\}$
RLSPCL_2	$F(\beta) \equiv \nabla(\frac{1}{2}\|H\beta - T\|_2^2)$, $G(\beta) \equiv \partial(\lambda\|\beta\|_2^2)$, $E = \{\beta \in \mathbb{R}^M : \|\beta\|_2^2 \leqslant \rho\}$

which typically indicates a T-score below −2.5 with a history of one or more fractures. Please note that the specific criteria for these stages may vary slightly depending on the guidelines and assessment methods used by healthcare professionals and organizations. It is essential to consult with a healthcare provider for a proper diagnosis and assessment of osteoporosis severity. The osteoporosis data used for each feature are detailed in Table 1.

We next explain the concept of an extreme learning machine (ELM) for applying our algorithms in machine learning data classification process. For N distinct samples, given a training set $u := \{(x^k, t^k) : x^k \in \mathbb{R}^n,\ t^k \in \mathbb{R}^m,\ k = 1, 2, \ldots, N\}$, where x^k denotes input training data and t^k represents the target. We focus on using single-hidden layer feed forward neural networks (SLFNs), the output function of ELM with M hidden nodes and sigmoid activation function can be represented by

$$O^k = \sum_{j=1}^{M} \beta^i \frac{1}{1 + \mathrm{e}^{-(w^j x^k + b^j)}}$$

where w^j and b^j are parameters of the beginning of random weight and bias, respectively. The finally optimal output weight β^i at the j-th hidden node by the hidden layer output matrix H is defined as follows:

$$H = \begin{bmatrix} \frac{1}{1+\mathrm{e}^{-(w^1 x^1 + b^1)}} & \cdots & \frac{1}{1+\mathrm{e}^{-(w^M x^1 + b^M)}} \\ \vdots & \ddots & \vdots \\ \frac{1}{1+\mathrm{e}^{-(w^1 x^N + b^1)}} & \cdots & \frac{1}{1+\mathrm{e}^{-(w^M x^N + b^M)}} \end{bmatrix}.$$

The goal of EML is to find optimal output weight $\beta = [\beta^{1^T}, \ldots, \beta^{M^T}]^T$ such that $T = H\beta$, where $T = [t^{1^T}, \ldots, t^{M^T}]^T$is the training target data matrix. To avoid over-fitting in the model. The least square regularization was considered in closed convex subsets as follows:

(i) Regularization of least square problem by L_1 (RLSPL_1) or the least absolute shrinkage and selection operator (LASSO): for $\lambda > 0$,

$$\min_{\beta \in R^M} \tfrac{1}{2}\|H\beta - T\|_2^2 + \lambda\|\beta\|_1. \tag{5}$$

(ii) Regularization of least square problem by L_2 (RLSPL_2): for $\lambda > 0$,

$$\min_{\beta \in R^M} \tfrac{1}{2}\|H\beta - T\|_2^2 + \lambda\|\beta\|_2^2. \tag{6}$$

(iii) Regularization of least square problem by L_1 with constrained by convex set L_1 (RLSPCL_1): for λ, $\rho > 0$,

$$\min_{\beta \in E} \tfrac{1}{2}\|H\beta - T\|_2^2 + \lambda\|\beta\|_1, \tag{7}$$

where $E = \{\beta \in \mathbb{R}^M : \|\beta\|_1 \leqslant \rho\}$.

(iv) Regularization of least square problem by L_2 with constrained by convex set L_2 (RLSPCL_2): for λ, $\rho > 0$,

$$\min_{\beta \in E} \tfrac{1}{2}\|H\beta - T\|_2^2 + \lambda\|\beta\|_2^2, \tag{8}$$

where $E = \{\beta \in \mathbb{R}^M : \|\beta\|_2^2 \leqslant \rho\}$.

For applying our algorithms to solve all of the convex minimization problems as above, we set our operator as in Table 2.

We considered four evaluation metrics, such as Accuracy, Precision, Recall, and F1-score, to evaluate the performance of the classification algorithms are defined as follows:

$$\begin{aligned} \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times 100\%, \\ \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\%, \\ \text{Recall} &= \frac{\text{TP}}{\text{TN} + \text{FN}} \times 100\%, \\ \text{F1-score} &= \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}, \end{aligned}$$

where these given matrices TP, TN, FP, and FN are the True Positive, True Negative, False Positives, and False Negatives, respectively.

The multi-class cross-entropy loss is a metric employed in classification tasks to assess how effectively a model distinguishes between multiple classes. This measurement is determined by computing the following average:

$$\text{Loss} = -\sum_{i=1}^{N} \mu_i \log \bar{\mu}_i,$$

where $\bar{\mu}_i$ represents the i-th scalar value in the model's output, μ_i stands for the corresponding target value for that specific scalar, and the variable N indicates the total number of scalar values in the entire model's output.

Next, we partition the dataset into 80% for training and 20% for testing. Our parameter settings include $\lambda^k = 1.99/\|H\|^2$, $\lambda = 10^{-5}$, and $M = 150$.

Table 3 Chosen parameters for algorithms.

Algorithm	α^k	σ^k	ε^k	ρ
IFBM	–	$\frac{2}{\|x^k-x^{k-1}\|^2+k^2}$	–	–
TPIPA	$\frac{0.7k}{k+1}$	$\frac{1}{\|x^k-x^{k-1}\|^3+k^3+2^3}$	$\frac{1}{\|x^k-x^{k-1}\|^5+k^5+2^5}$	–
Algorithm 1 (RLSPL_1)	$\frac{0.9k}{k+1}$	$\frac{2^{15}}{\|x^k-x^{k-1}\|^3+k^3+2^{15}}$	$\frac{1}{100k+1}$	–
Algorithm 1 (RLSPL_2)	$\frac{0.5k}{k+1}$	$\frac{2^{12}}{\|x^k-x^{k-1}\|^2+k^2+2^{12}}$	$\frac{1}{10k+1}$	–
Algorithm 1 (RLSPCL_1)	$\frac{k}{5k+1}$	$\frac{2^{15}}{\|x^k-x^{k-1}\|^2+k^2+2^{15}}$	$\frac{1}{10k+1}$	4
Algorithm 1 (RLSPCL_2)	$\frac{0.9k}{k+1}$	$\frac{2^{10}}{\|x^k-x^{k-1}\|^2+k^2+2^{10}}$	$\frac{1}{\|x^k-x^{k-1}\|^3+k^3+2^3}$	9

Table 4 The performance using Precision, Recall, F1-score, and Accuracy.

Algorithm	Iteration no.	Training time	Precision	Recall	F1-score	Accuracy
IFBM	480	0.0074	81.44	76.94	79.13	80.37
TPIPA	468	0.0094	82.76	75.82	79.14	80.25
Algorithm 1 (RLSPL_1)	280	0.0101	83.33	77.21	80.15	80.82
Algorithm 1 (RLSPL_2)	382	0.0113	82.86	77.13	79.89	80.71
Algorithm 1 (RLSPCL_1)	300	0.0084	81.99	76.96	79.40	80.48
Algorithm 1 (RLSPCL_2)	409	0.0078	82.65	77.19	79.82	80.71

We evaluate and compare the performance of IFBM, TPIPA, and our algorithms with a complete list of parameters given in Table 3 when

$$\theta^k = \begin{cases} \frac{1}{\|x^k-x^{k-1}\|k^3}, & \text{if } x^k \neq x^{k-1} \text{ and } k > N, \\ \sigma^k, & \text{otherwise,} \end{cases}$$

and

$$\delta^k = \begin{cases} \frac{1}{\|x^{k-1}-x^{k-2}\|k^3}, & \text{if } x^{k-1} \neq x^{k-2} \text{ and } k > N, \\ \varepsilon^k, & \text{otherwise,} \end{cases}$$

where N is the iteration number that we want to stop.

Table 4 shows that our algorithm exhibits the highest efficiency in terms of precision, recall, F1-score, and accuracy efficiency. It demonstrates the highest likelihood of correctly classifying osteoporosis compared to the above mentioned algorithms.

Next, we compare our method with machine learning methods in terms of accuracy using the same set of information. The results are presented in Table 5.

In Table 5, we see that the method studied is the highest efficient in accuracy, precision, recall, and F1-score, establishing it as the most accurate predictor of osteoporosis. Next, we present accuracy and loss graphs for both the training and testing data to assess the potential over-fitting of our algorithm.

From Figs. 1–4, we can see that training loss and validation loss values tend to decrease until a certain point, after which they stabilize. In contrast, when we assess the accuracy graph, it becomes apparent that both training accuracy and validation accuracy exhibit an upward trend, with validation accuracy consistently surpassing training accuracy.

CONCLUSION

In this study, we introduce a projective bi-inertial forward-backward splitting algorithm for solving the sum of two monotone operators in a real Hilbert space;

Table 5 The highest accuracy of different machine learning methods using bone mineral density dataset.

Algorithm	Validation	Accuracy
SVM kernel	10-fold cross validation	77.60
Boosted Trees	5-fold cross validation	60.40
Logistic Regression kernel	5-fold cross validation	68.60
Medium kNN	5-fold cross validation	79.80
Algorithm 1 (RLSPL_1)	Training data(80%), Validation (20%)	80.82
Algorithm 1 (RLSPL_2)	Training data (80%), Validation (20%)	80.71
Algorithm 1 (RLSPCL_1)	Training data (80%), Validation (20%)	80.48
Algorithm 1 (RLSPCL_2)	Training data (80%), Validation (20%)	80.71

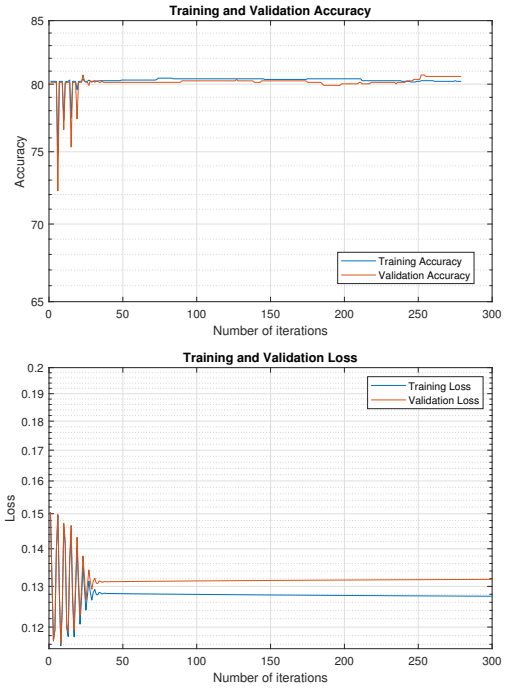

Fig. 1 Accuracy and Loss plots of training and validation for the iterations of Algorithm 1 (RSLPL_1).

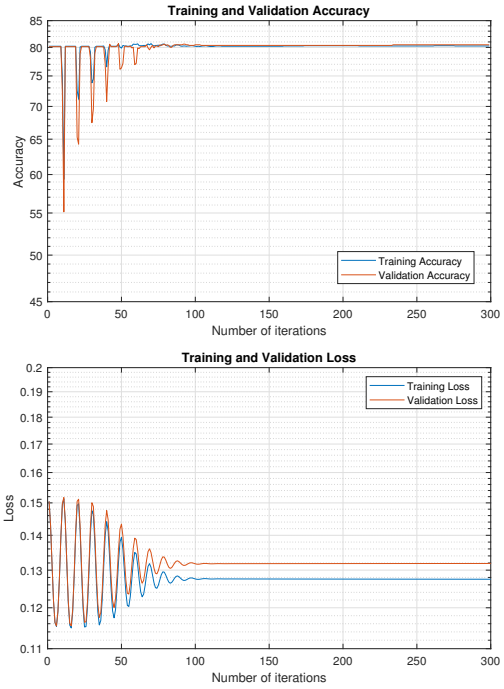

Fig. 3 Accuracy and Loss plots of training and validation for the iterations of Algorithm 1 (RSLPCL_1).

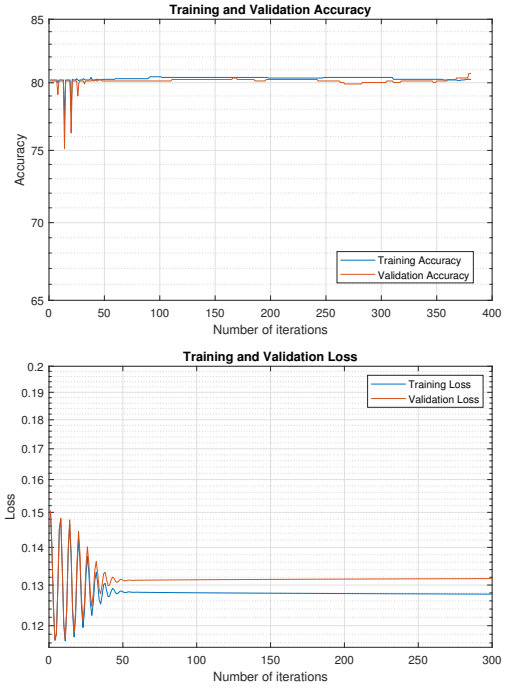

Fig. 2 Accuracy and Loss plots of training and validation for the iterations of Algorithm 1 (RSLPL_2).

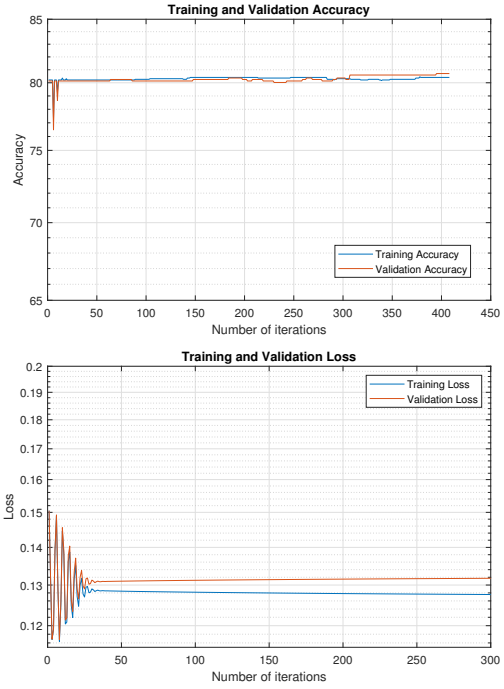

Fig. 4 Accuracy and Loss plots of training and validation for the iterations of Algorithm 1 (RSLPCL_2).

one is maximally monotone, and the other is Lipschitz continuous. We establish a weak convergence theorem for the sequence generated by our algorithm, subject to appropriate conditions. Furthermore, we utilize our algorithm as a machine learning tool within the context of the extreme learning machine model for a classification problem. We employ a reliable dataset for predicting bone mineral density to assess the algorithm's effectiveness. The study results indicate that our algorithm is more efficient than the machine learning methods in Table 5.

Data Availability

The dataset used in this research is publicly available at: https://dataverse.harvard.edu/file.xhtml?fileId=6563909&version=1.0.

Acknowledgements: This research was supported by the National Research Council of Thailand (N42A650334) and Thailand Science Research and Innovation, University of Phayao (Fundamental Fund 2024).

REFERENCES

1. Chen P, Huang J, Zhang X (2013) A primal-dual fixed point algorithm for convex separable minimization with applications to image restoration. *Inverse Probl* **29**, 025011.
2. Combettes PL, Wajs VR (2005) Signal recovery by proximal forward-backward splitting. *Multiscale Model Simul* **4**, 1168–1200.
3. Padcharoen A, Kitkuan D, Kumam W, Kumam P (2021) Tseng methods with inertial for solving inclusion problems and application to image deblurring and image recovery problems. *Comput Math Methods* **3**, e1088.
4. Byrne C (2003) A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse Probl* **20**, 103.
5. Gibali A, Thong DV (2018) Tseng type methods for solving inclusion problems and its applications. *Calcolo* **55**, 49.
6. Hanjing A, Suantai S (2020) A fast image restoration algorithm based on a fixed point and optimization method. *Mathematics* **8**, 378.
7. Thong DV, Cholamjiak P (2019) Strong convergence of a forward-backward splitting method with a new step size for solving monotone inclusions. *Comput Appl Math* **38**, 94.
8. Lions PL, Mercier B (1979) Splitting algorithms for the sum of two nonlinear operators. *SIAM J Numer Anal* **16**, 964–979.
9. Passty GB (1979) Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J Math Anal Appl* **72**, 383–390.
10. Tseng P (2000) A modified forward-backward splitting method for maximal monotone mappings. *SIAM J Control Optim* **38**, 431–446.
11. Bauschke HH, Combettes PL (2011) Convexity and nonexpansiveness. In: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, *CMS Books in Mathematics*, Springer, New York, pp 59–74.
12. Eckstein J, Bertsekas DP (1992) On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math Program* **55**, 293–318.
13. Abubakar J, Kumam P, Rehman HU, Hassan Ibrahim A (2020) Inertial iterative schemes with variable step sizes for variational inequality problem involving pseudomonotone operator. *Mathematics* **8**, 609.
14. Alvarez F (2000) On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM J Control Optim* **38**, 1102–1119.
15. Alvarez F, Attouch H (2001) An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Anal* **9**, 3–11.
16. Nesterov Y (1983) A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. *Dokl Akad Nauk SSSR* **269**, 543.
17. Ceng LC, Petruşel A, Qin X, Yao JC (2020) A Modified inertial subgradient extragradient method for solving pseudomonotone variational inequalities and common fixed point problems. *Fixed Point Theory* **21**, 93–108.
18. He L, Cui YL, Ceng LC, Zhao TY, Wang DQ, Hu HY (2021) Strong convergence for monotone bilevel equilibria with constraints of variational inequalities and fixed points using subgradient extragradient implicit rule. *J Inequalities Appl* **2021**, 146.
19. Zhao TY, Wang DQ, Ceng LC, He L, Wang CY, Fan HL (2020) Quasi-inertial Tseng's extragradient algorithms for pseudomonotone variational inequalities and fixed point problems of quasi-nonexpansive operators. *Numer Funct Anal Optim* **42**, 69–90.
20. Iyiola OS, Shehu Y (2022) Convergence results of two-step inertial proximal point algorithm. *Appl Numer Math* **182**, 57–75.
21. Goebel K, Kirk WA (1990) *Topics in Metric Fixed Point Theory*, Cambridge University Press, Cambridge.
22. López G, Martín-Márquez V, Wang F, Xu HK (2012) Forward-backward splitting methods for accretive operators in Banach spaces. *Abstr Appl Anal* **2012**, 109236.
23. Opial Z (1967) Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull Amer Math Soc* **73**, 591–597.
24. Ofoedu EU (2006) Strong convergence theorem for uniformly L-Lipschitzian asymptotically pseudocontractive mapping in real Banach space. *J Math Anal Appl* **321**, 722–728.
25. Yao Q, Lin J, Wei S (2023) H_2O_2 suppresses ROS elimination and osteogenesis by inhibiting MT2 expression. *ScienceAsia* **49**, 70–76.
26. Linfeng H (2022) Harvard Dataverse: Bone mineral density, avaiable at: https://doi.org/10.7910/DVN/UDZIJS.