J. Math. Computer Sci., 37 (2025), 190–200

Online: ISSN 2008-949X

Journal of Mathematics and Computer Science

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

Screening osteoporosis in elderly using a new two inertial projective forward-backward splitting algorithm

Chakarach Mekkrua[a], Pronpat Peeyada[b], Watcharaporn Cholamjiak[b,*], Wongthawat Liawrungrueang[c]

[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.

Abstract

In this work, we propose a new two inertial projective forward-backward splitting algorithm for approximating the solution of the variational inclusion problem in real Hilbert spaces. We prove weak convergence of the sequence generated by our proposed iterative algorithm. Moreover, we also provide an application to predict osteoporosis in the elderly using a dataset from the Harvard Dataverse. The comparison of algorithm performance is calculated using accuracy, precision, recall, and F1-score. Our algorithm's performance is higher than other comparable algorithms. As a result, our algorithm is an effective classification technique for identifying osteoporosis.

Keywords: Variational inclusion problem, osteoporosis, elderly, inertial method, data classification.

2020 MSC: 46E20, 52A07, 68Q04.

1. Introduction

In this article, we consistently assume that $\mathcal{H}$ is a real Hilbert space equipped with the norm $\|.\|$ and the inner product $\langle .,. \rangle$, $F : \mathcal{H} \to \mathcal{H}$ is a single-valued monotone operator, and $G : \mathcal{H} \to 2^{\mathcal{H}}$ is a set-valued monotone operator. We know that the variational inclusion problem (VIP) can be formulated as the following problem:

$$\text{Find an element } u \in \mathcal{H} \text{ such that } 0 \in (Fu + Gu). \tag{1.1}$$

The variational inclusion problem has garnered significant attention due to its central role in several fundamental concepts in applied mathematics, including convex minimization, split feasibility, fixed-point, saddle point, variational inequality, and equilibrium problems (see [3, 4, 6, 19, 20]). Moreover, it serves as a model for various problems in applied sciences and engineering disciplines, such as signal processing, image reconstruction, approximation theory, control theory, biomedical engineering, communications, and geophysics. For additional references, one can consult [2, 5, 10, 11, 14, 18] and related literature.

*Corresponding author

Email address: watcharaporn.ch@up.ac.th (Watcharaporn Cholamjiak)

doi: 10.22436/jmcs.037.02.04

Received: 2024-02-27 *Revised:* 2024-07-21 *Accepted:* 2024-08-01

One of the earliest methods for solving equation (1.1) is the forward-backward splitting method defined as follows:

$$u^{k+1} = J_\lambda^G(u^k - \lambda F u^k),\ \forall k \in \mathbb{N}, \tag{1.2}$$

where $J_\lambda^G = (I + \lambda G)^{-1}$ is the resolvent of the operator G and $\lambda \in (0, \frac{2}{L})$. The sequences generated by algorithm (1.2) demonstrate weak convergence towards a solution of (1.1) when G is $\frac{1}{L}$-inverse strongly monotone (or cocoercive). Another viable condition for convergence of (1.2) is requiring similar strong monotonicity for $F + G$.

To solve the zero-finding problem for the sum of two monotone operators, the inertial proximal algorithm is used. This algorithm incorporates the inertial technique into the forward-backward algorithm, resulting in the inertial forward-backward algorithm (IFBA), as proposed by Moudafi and Oliny [15]. Let x^0 and x^1 be in H, and let λ^k be in the interval $(0, \frac{2}{L})$, where L is the Lipschitz constant of F, for all $k \geqslant 1$:

$$\begin{cases} v^k = u^k + \theta^k(u^k - u^{k-1}), \\ u^{k+1} = J_{\lambda^k}^G(v^k - \lambda^k F u^k),\ k \geqslant 1. \end{cases}$$

They established the weak convergence of the iterative sequence by establishing conditions relying on u^k and the parameter θ^k, which are contingent on the cocoercivity assumption regarding F and the solution set. In 2015, Lorenz and Pock [13] introduced a modification to the inertial forward-backward splitting algorithm (IFBSA). Their algorithm is defined as follows:

$$\begin{cases} v^k = u^k + \theta^k(u^k - u^{k-1}), \\ u^{k+1} = J_{\lambda^k}^G(v^k - \lambda^k F v^k),\ k \geqslant 1, \end{cases}$$

where $\theta^k \in [0, 1)$ is an extrapolation factor and λ^k is a step size parameter in positive real interval. They proved that the iterative sequence generated by IFBSA converges weakly to a zero of the sum of two maximal monotone operators F and G.

Iyiola and Shehu [9] introduced and studied the two-point inertial proximal point algorithm (TPIPA) for monotone operators in Hilbert spaces. They employed the following iterative algorithm to prove a weak convergence result:

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

where $\lambda^k > 0$, θ^k is in $[0, \infty)$, and δ^k is relaxed conditions in $(-\infty, 0]$.

Motivated by the abovementioned results, we introduce a new two inertial projective forward-backward splitting algorithm to approximate the solution of the variational inclusion problem in real Hilbert spaces. Furthermore, we provide a proof of weak convergence and discuss some consequences of our main result. We also present an application for predicting osteoporosis and compare the performance of our algorithm with the literature mentioned earlier in terms of accuracy, precision, recall, and F1- score.

2. Preliminaries

In this section, we denote the weak and strong convergence of a sequence u^k to a point $u \in \mathcal{H}$ as $u^k \rightharpoonup u$ and $u^k \to u$, respectively. We are listing some crucial results necessary to prove our main result.

Definition 2.1. An operator $F : \mathcal{H} \to \mathcal{H}$ is said to be:

(i) monotone if $\langle Fu - Fv, u - v\rangle \geqslant 0,\ \forall u, v \in \mathcal{H}$;

(ii) L-Lipschitz continuous if there is a constant $L > 0$ such that $\|Fu - Fv\| \leqslant L\|u - v\|,\ \forall u, v \in \mathcal{H}$, if $L = 1$, then F is said to be nonexpansive;

(iii) firmly nonexpansive if

$$\|Fu - Fv\|^2 \leqslant \|u - v\|^2 - \|(I - F)u - (I - F)v\|^2, \ \forall u, v \in \mathcal{H},$$

or equivalently

$$\langle Fu - Fv, u - v\rangle \geqslant \|Fu - Fv\|^2, \ \forall u, v \in \mathcal{H};$$

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

Lemma 2.2 ([7]). *Let $F : \mathcal{H} \rightharpoonup \mathcal{H}$ be a nonexpansive mapping such that $\mathrm{Fix}(F) \neq \Phi$. If there exists a sequence $\{u^k\}$ in $\mathcal{H}$ such that $u^k \rightharpoonup u \in H$ and $\|u^k - Fu^k\| \rightharpoonup 0$, then $u \in \mathrm{Fix}(F)$.*

Lemma 2.3 ([12]). *Let $F : \mathcal{H} \to \mathcal{H}$ be a τ-cocoercive mapping and $G : \mathcal{H} \to 2^{\mathcal{H}}$ be a maximal monotone mapping. Then, we have*

(i) *for $\lambda > 0$, $\mathrm{Fix}(J^G_\lambda(I - \lambda F)) = (F + G)^{-1}(0)$;*
(ii) *for $0 < \lambda < \overline{\lambda}$ and $x \in H$, $\|u - J^G_\lambda(I - \lambda F)u\| \leqslant 2\|u - J^G_{\overline{\lambda}}(I - \overline{\lambda}F)u\|$.*

Lemma 2.4 ([17]). *Let Ω be a nonempty set of $\mathcal{H}$ and $\{u^k\}$ be a sequence in $\mathcal{H}$. Assume that the following conditions hold.*

(i) *For every $u \in \Omega$, the sequence $\{\|u^k - u\|\}$ converges.*
(ii) *Every weak sequential cluster point of $\{u^k\}$ belongs to Ω.*

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

Lemma 2.5 ([16]). *Suppose that $\{\gamma^k\}$, $\{\xi^k\}$ and $\{\upsilon^k\}$ are sequences in $[0, +\infty)$ such that $\gamma^{k+1} \leqslant \gamma^k + \upsilon^k(\gamma^k - \gamma^{k-1}) + \xi^k$, $\forall k \geqslant 1$, $\sum_{k=1}^{\infty} \xi^k < +\infty$ and there is $\upsilon \in \mathbb{R}$ with $0 \leqslant \upsilon^k < \upsilon < 1$, $\forall k \geqslant 1$. Then the following conditions are satisfied:*

(i) *$\sum[\gamma^k - \gamma^{k-1}]_+ < +\infty$, where $[r]_+ = \max\{r, 0\}$;*
(ii) *there exists $\gamma^* \in [0, \infty)$ such that $\lim_{k\to+\infty} \gamma^k = \gamma^*$.*

3. Main results

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

Algorithm 3.1. Two inertial projective forward-backward splitting algorithm for (VIP).
Initialization: Select arbitrary points $u^{-1}, u^0, v^0 \in \mathcal{H}, \{\alpha^k\} \subset (a, b) \subset (0, 1], \{\lambda^k\} \subset (c, d) \subset (0, 2\tau)$, and $\{\theta^k\}, \{\delta^k\} \subset (-\infty, \infty)$. Set $k = 0$
Iterative Steps: Compute v^{k+1} as follows:

Step 1. Calculate $u^{k+1} = (1 - \alpha^k)v^k + \alpha^k J^k v^k$.
Step 2. Calculate $v^{k+1} = P_E(u^{k+1} + \theta^k(u^{k+1} - u^k) + \delta^k(u^k - u^{k-1}))$, where $J^k = J^k_{\lambda^k}(I - \lambda^k F)$. Set $k := k + 1$ and go to Step 1.

Theorem 3.2. *Let $\{v^k\}$ be a sequence generated by Algorithm 3.1 which satisfies the following conditions:*

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

Then $\{v^k\}$ converges weakly to a solution of $(F + G)^{-1}(0) \cap E$.

Proof. Let $p \in (F+G)^{-1}(0) \cap E$. Since J^k is nonexpansive when $\{\lambda^k\} \subset (0, 2\tau)$ and the projection mapping P_E is also nonexpansive, we have

$$\begin{aligned}
\|v^{k+1} - p\| &= \|P_E(u^{k+1} + \theta^k(u^{k+1} - u^k) + \delta^k(u^k - u^{k-1})) - p\| \\
&\leqslant \|u^{k+1} - p\| + \theta^k\|u^{k+1} - u^k\| + \delta^k\|u^k - u^{k-1}\| \\
&\leqslant (1-\alpha^k)\|v^k - p\| + \alpha^k\|J^k v^k - p\| + \theta^k((1-\alpha^k)\|v^k - u^k\| \\
&\quad + \alpha^k\|J^k v^k - u^k\|) + \delta^k\|u^k - u^{k-1}\| \\
&\leqslant (1-\alpha^k)\|v^k - p\| + \alpha^k\|v^k - p\| + \theta^k((1-\alpha^k)\|v^k - u^k\| + \alpha^k\|v^k - u^k\|) + \delta^k\|u^k - u^{k-1}\| \\
&= \|v^k - p\| + |\theta^k|\|v^k - u^k\| + |\delta^k|\|u^k - u^{k-1}\|.
\end{aligned}$$

By condition (i), it follows from Lemma 2.5 that $\lim_{k\to\infty}\|v^k - p\|$ exists. This implies that $\{u^k\}$ is bounded. Since $J^G_{\lambda^k}$ is firmly nonexpensive and P_E is nonexpansive, we have

$$\begin{aligned}
\|v^{k+1} - p\|^2 &= \|P_E(u^{k+1} + \theta^k(u^{k+1} - u^k) + \delta^k(u^k - u^{k-1})) - p\|^2 \\
&\leqslant \|u^{k+1} - p\|^2 + 2\langle\theta^k(u^{k+1} - u^k) + \delta^k(u^k - u^{k-1}), u^{k+1} + \theta^k(u^{k+1} - u^k) \\
&\quad + \delta^k(u^k - u^{k-1}) - p\rangle \\
&\leqslant (1-\alpha^k)\|v^k - p\|^2 + \alpha^k\|J^k v^k - p\|^2 + 2\langle\theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}), v^k \\
&\quad + \theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}) - p\rangle \\
&\leqslant (1-\alpha^k)\|v^k - p\|^2 + \alpha^k[\|v^k - \lambda^k Fv^k - p + \lambda^k Fp\|^2 \\
&\quad - \|v^k - \lambda^k Fv^k - J^k v^k - p + \lambda^k Fp + J^k p\|^2] + 2\langle\theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}), v^k \\
&\quad + \theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}) - p\rangle \\
&= (1-\alpha^k)\|v^k - p\|^2 + \alpha^k[\|v^k - p\|^2 - 2\lambda^k\langle v^k - p, Fv^k - Fp\rangle + (\lambda^k)^2\|Fv^k - Fp\|^2] \\
&\quad - \alpha^k\|v^k - \lambda^k(Fv^k - Fp) - J^k v^k\|^2 + 2\langle\theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}), v^k \\
&\quad + \theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}) - p\rangle \\
&\leqslant \|v^k - p\|^2 + \alpha^k(\lambda^k)^2\|Fv^k - Fp\|^2 - 2\tau\alpha^k\lambda^k\|Fv^k - Fp\|^2 \\
&\quad - \alpha^k\|v^k - \lambda^k(Fv^k - Fp) - J^k v^k\|^2 + 2\langle\theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}), v^k \\
&\quad + \theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}) - p\rangle \\
&= \|v^k - p\|^2 - \alpha^k\lambda^k(2\tau - \lambda^k)\|Fv^k - Fp\|^2 - \alpha^k\|v^k - \lambda^k(Fv^k - Fp) - J^k v^k\|^2 \\
&\quad + 2\langle\theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}), v^k + \theta^k(v^k - u^k) + \delta^k(u^k - u^{k-1}) - p\rangle.
\end{aligned} \tag{3.1}$$

Again by the conditions (i)-(iii) and (3.1), we have

$$\lim_{k\to\infty}\|Fv^k - Fp\| = \lim_{k\to\infty}\|v^k - \lambda^k(Fv^k - Fp) - J^k v^k\| = 0.$$

This implies that $\lim_{k\to\infty}\|v^k - J^k v^k\| = 0$. Since $\liminf_{k\to\infty}\lambda^k > 0$, there is $\lambda > 0$ such that $\lambda^k > \lambda$. By Lemma 2.3 (ii), we obtain

$$\lim_{k\to\infty}\|v^k - J^G_\lambda(I - \lambda F)v^k\| \leqslant \lim_{k\to\infty}\|v^k - J^k v^k\| = 0. \tag{3.2}$$

Since $\{v^k\}$ is bounded, we can let $\widehat{v}$ be a weak sequential cluster point of $\{v^k\}$. By applying Lemma 2.2 and (3.2), we can get that $\widehat{v} \in \mathrm{Fix}(J^G_\lambda(I - \lambda F)) = (F+G)^{-1}(0)$. Since v^k is a sequence in E and E is closed, it follows that $\widehat{v} \in (F+G)^{-1}(0) \cap E$. By utilizing opial's lemma (Lemma 2.4), we can obtain that $\widehat{v}$ weakly converges to an element in $(F+G)^{-1}(0) \cap E$. □

4. Application

Currently, osteoporosis is the most common bone-weakening disease characterized by decreased bone mass and deteriorated bone tissue, leading to an increased risk of fractures, especially in the elderly population. According to the International Osteoporosis Foundation (IOF), more than 200 million people worldwide suffer from osteoporosis, with 1 in 3 women and 1 in 5 men over the age of 50 years being affected. In Thailand, the proportion of elderly individuals is steadily rising. Over 1 million people are living with osteoporosis, and shockingly, 1 in 4 Thais are unaware that osteoporosis can lead to paralysis and death in its final stage. A first-time fracture due to osteoporosis often leads to second and third fractures. Early screening of this disease in elderly patients is crucial to preventing undesirable outcomes. Traditional methods of osteoporosis diagnosis, such as dual-energy X-ray absorptiometry measurements, have limitations in terms of accessibility and cost-effectiveness, especially in developing countries like Thailand. Due to these challenges, the opportunity to utilize this approach remains inadequate. To screen patients quickly and accurately, we used the elderly osteoporosis dataset from the Harvard Dataverse available on internet website [8]. We employed 1,159 records and 37 attributes and one class to evaluate the proposed algorithm for training and testing.

Osteoporosis is typically categorized into four stages or levels based on bone density measurements using a standard called T-scores. These levels are as follows.

(i) Normal: In this stage, bone density is considered normal, and the T-score is above -1.0.
(ii) Osteopenia: Osteopenia is a condition characterized by lower bone density than normal but not low enough to be classified as osteoporosis. T-scores in this stage typically range from -1.0 to -2.5.
(iii) Osteoporosis: This is the advanced stage of bone loss where bone density is significantly reduced. T-scores in this stage are typically -2.5 or lower.
(iv) Severe Osteoporosis: In some classifications, there may be a further stage called "severe osteoporosis," which typically indicates a T-score below -2.5 with a history of one or more fractures.

It's important to note that these stages are determined by bone density measurements and don't necessarily correlate with the presence or absence of symptoms. A person with osteoporosis or severe osteoporosis may not have any noticeable symptoms until they experience a fracture or other complications. Diagnosis and management of osteoporosis are typically done by healthcare professionals based on bone density measurements, clinical assessment, and individual risk factors. The osteoporosis data associated with each feature is detailed in Table 1.

In the process of machine learning data classification, we focused on an extreme learning machine (ELM), which was defined as follows. Assume we have N distinct samples such that $u := \{(u^k, t^k) : u^k \in R^n, t^k \in R^m, k = 1, 2, \ldots, N\}$ is a set of training data, where u^k and t^k are an input training data and training target target, respectively. The output function of ELM for a standard single hidden layer feedforward networks (SLFNs) with M hidden nodes are mathematically modeled as:

$$O^j = \sum_{i=1}^{M} \mu^i \frac{1}{1 + e^{-(w^i u^j + b^i)}},$$

where A activation function with w^i is parameter weight and b^i is bias. To find optimal output weight μ^i at the i-th hidden node, the above equations can be denoted in the form of matrix as $T = \mathcal{A}\mu$, where

$$\mathcal{A} = \begin{bmatrix} \frac{1}{1+e^{-(w^1 u^1 + b^1)}} & \cdots & \frac{1}{1+e^{-(w^M u^1 + b^M)}} \\ \vdots & \ddots & \vdots \\ \frac{1}{1+e^{-(w^1 u^N + b^1)}} & \cdots & \frac{1}{1+e^{-(w^M u^N + b^M)}} \end{bmatrix},$$

$\mathcal{A}$ is called the hidden layer output matrix, $T = [t^{1^T}, \ldots, t^{M^T}]^T$ represents the training target data matrix, and $\mu = [\mu^{1^T}, \ldots, \mu^{M^T}]^T$ represents optimal output weight such that $\mu = \mathcal{A}^\dagger T$ where $\mathcal{A}^\dagger$ is the Moore-

Table 1: Overview of osteoporosis dataset.

Attribute Name	Max	Min	Mean	Median	Standard Deviation
Gender	2	1	1.3794	1	0.48545
Age	99.8	50	64.0635	61.9	00.053
Height	186	141	165.3392	166	8.2434
Weight	113	23	66.9667	67	11.7686
BMI	36.7537	9.2133	24.3976	24.3375	3.2862
L1.4T	6	-3.4	-0.63862	-0.8	1.5679
FNT	2.7	-5.05	-1.3004	-1.4	1.1182
TLT	3	-4.65	-0.92414	-1	1.1563
ALT	181	4	23.1963	19	16.0187
AST	128	10	22.4322	21	9.4318
BUN	69.8	1.79	5.5995	5.18	3.358
CREA	381.2	5.86	74.0888	70.5	27.1017
URIC	745.3	5.46	348.8885	340.8	98.227
FBG	21.67	3.13	5.3225	4.95	1.5172
HDL-C	5.46	0.45	1.2491	1.19	0.38531
LDL-C	6.65	0.14	2.5875	2.54	0.88982
Ca	5.84	1.78	2.2347	2.23	0.17088
P	4.41	0.56	1.0377	1.02	0.21513
Mg	1.73	0.097	0.86274	0.86	0.10066
Calsium	1	0	0.15039	0	0.35761
Calcitriol	1	0	0.16595	0	0.37219
Bisphosph	1	0	0.061366	0	0.2401
Calcitonin	1	0	0.061366	0	0.2401
HTN	1	0	0.54797	1	0.49791
COPD	1	0	0.24201	0	0.42848
DM	1	0	0.33016	0	0.47048
Hyperlipid	1	0	0.39153	0	0.4883
Hyperuricemia	1	0	0.17805	0	0.38272
AS	1	0	0.75194	1	0.43207
VT	1	0	0.01815	0	0.13355
VD	1	0	0.075194	0	0.26382
OP	1	0	0.37597	0	0.48458
CAD	1	0	0.21089	0	0.40812
CKD	1	0	0.038029	0	0.19135
Smoking	1	0	0.25151	0	0.43407
Drinking	1	0	0.22126	0	0.41528
Fracture	1	0	0.015557	0	0.1281

Penrose generalized inverse of $\mathcal{A}$, it may be difficult to find when the matrix $\mathcal{A}$ does not exist. Thus, finding such a solution μ through convex minimization can overcome such difficulty.

We consider regularization of least square problems using techniques like L_1 (Lasso) and L_2 (Ridge) regularization. Regularization is a technique commonly used in machine learning and statistics to prevent overfitting and improve the generalization of models, which can lead to better generalization in classification problems. We conduct a series of experiments on a classification problem. The specific details of these problems are provided below.

(i) Regularization of least square problem by L_1 (RLSP-L_1) or well-known called the least absolute shrink-

age and selection operator (LASSO): for $\lambda > 0$,

$$\min_{\mu \in R^M} \frac{1}{2}\|\mathcal{A}\mu - T\|_2^2 + \lambda\|\mu\|_1. \tag{4.1}$$

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

$$\min_{\mu \in R^M} \frac{1}{2}\|\mathcal{A}\mu - T\|_2^2 + \lambda\|\mu\|_2^2. \tag{4.2}$$

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

$$\min_{\mu \in E} \frac{1}{2}\|\mathcal{A}\mu - T\|_2^2 + \lambda\|\mu\|_1, \tag{4.3}$$

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

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

$$\min_{\mu \in E} \frac{1}{2}\|\mathcal{A}\mu - T\|_2^2 + \lambda\|\mu\|_2^2, \tag{4.4}$$

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

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

Table 2: Setting operators of our algorithms to solve all of the convex minimization problems (4.1)-(4.4).

Problem	**Setting operator of our algorithms**
RLSP-L_1	$F(\mu) \equiv \nabla(\frac{1}{2}\|\mathcal{A}\mu - T\|_2^2)$, $G(\mu) \equiv \partial(\lambda\|\mu\|_1)$, $E = \mathcal{H}$
RLSP-L_2	$F(\mu) \equiv \nabla(\frac{1}{2}\|\mathcal{A}\mu - T\|_2^2)$, $G(\mu) \equiv \partial(\lambda\|\mu\|_2^2)$, $E = \mathcal{H}$
RLSPC-L_1	$F(\mu) \equiv \nabla(\frac{1}{2}\|\mathcal{A}\mu - T\|_2^2)$, $G(\mu) \equiv \partial(\lambda\|\mu\|_1)$, $E = \{\mu \in R^M : \|\mu\|_1 \leqslant \rho\}$
RLSPC-L_2	$F(\mu) \equiv \nabla(\frac{1}{2}\|\mathcal{A}\mu - T\|_2^2)$, $G(\mu) \equiv \partial(\lambda\|\mu\|_2^2)$, $E = \{\mu \in R^M : \|\mu\|_2^2 \leqslant \rho\}$

We examined four assessment criteria, including accuracy, precision, recall, and F1-score [21], to assess the effectiveness of the classification algorithms. These metrics are defined as follows:

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

In these matrices, TP is the True Positive, TN is the True Negative, FP is the False Positive, and FN is the False Negative.

The multi-class cross-entropy loss [1] is a metric used in classification tasks to evaluate a model's ability to distinguish between multiple classes. This measurement is determined by calculating the following average:

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

where $\bar{\varphi}_i$ denotes the i-th scalar value in the model's output, while φ_i signifies the corresponding target value associated with that particular scalar. The variable N represents the total count of scalar values encompassed within the entire model's output.

Next, we partition the dataset into 70% for training and 30% for testing. Our parameter settings consist of $\lambda^k = \frac{1.99}{\|\mathcal{H}\|^2}$, $\lambda = 10^{-5}$, and $M = 170$. We assess and compare the performance of IFBA, IFBSA, TPIPA, and our algorithm with a comprehensive set of parameters listed in Table 3, where

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

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

Table 3: Chosen parameters of each algorithm.

Algorithm	α^k	σ^k	ε^k	ρ
IFBA	-	$\frac{10^{10}}{\|x^k - x^{k-1}\|^3 + k^3 + 10^{10}}$	-	-
IFBSA	-	$\frac{1}{2k+1}$	-	-
TPIPA	$\frac{k}{2k+1}$	$\frac{1}{100k+1}$	$\frac{1}{2k+1}$	-
Algorithm 3.1 (RLSP-L_1)	$\frac{0.9k}{k+1}$	$\frac{1}{\|x^k - x^{k-1}\|^3 + k^3 + 2^3}$	$\frac{2^{12}}{\|x^k - x^{k-1}\|^3 + k^3 + 2^{12}}$	-
Algorithm 3.1 (RLSP-L_2)	$\frac{0.7k}{k+1}$	$\frac{2^{15}}{\|x^k - x^{k-1}\|^3 + k^3 + 2^{15}}$	$\frac{1}{2k+1}$	-
Algorithm 3.1 (RLSPC-L_1)	$\frac{0.9k}{k+1}$	$\frac{2^{10}}{\|x^k - x^{k-1}\|^3 + k^3 + 2^{10}}$	$\frac{1}{\|x^k - x^{k-1}\|^3 + k^3 + 2^3}$	4
Algorithm 3.1 (RLSPC-L_2)	$\frac{0.9k}{k+1}$	$\frac{1}{100k+1}$	$\frac{2^{12}}{\|x^k - x^{k-1}\|^3 + k^3 + 2^{12}}$	4

Table 4: Comparison of the performance with each algorithm.

Algorithm	**Iteration No.**	**Training Time**	**Precision**	**Recall**	**F1-score**	**Accuracy**
IFBA	681	0.0092	78.88	77.26	78.06	78.30
IFBSA	700	0.0082	79.86	76.95	78.38	78.88
TPIPA	949	0.0115	81.74	75.68	78.59	78.88
Algorithm 3.1 (RLSP-L_1)	662	0.0082	81.52	76.29	78.82	79.02
Algorithm 3.1 (RLSP-L_2)	572	0.0085	81.22	76.35	78.71	79.02
Algorithm 3.1 (RLSPC-L_1)	674	0.0095	81.52	76.29	78.82	79.02
Algorithm 3.1 (RLSPC-L_2)	609	0.0086	81.22	76.35	78.71	79.02

Table 4 demonstrates our algorithm's exceptional precision, recall, F1-score, and accuracy performance. It also has the lowest number of iterations. This indicates its superior efficiency, making it the most likely candidate for accurate osteoporosis classification among the 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.

Table 5: Highest accuracy of different machine learning methods using bone mineral density dataset.

Algorithm	**Validation**	**Accuracy**
Kernel	k-fold cross validation, $k = 5$	67.20
Ensemble	k-fold cross validation, $k = 5$	57.50
k-nearest neighbors (kNN)	k-fold cross validation, $k = 10$	77.20
Algorithm 3.1 (RLSP-L_1)	Train (70%), test (30%)	79.02
Algorithm 3.1 (RLSP-L_2)	Train (70%), test (30%)	79.02
Algorithm 3.1 (RLSPC-L_1)	Train (70%), test (30%)	79.02
Algorithm 3.1 (RLSPC-L_2)	Train (70%), test (30%)	79.02

Table 5 shows that the method studied is the most efficient in accuracy compared to machine learning methods, thus establishing it as the most accurate predictor of osteoporosis. Next, we will present accuracy and loss graphs for both the training and testing data to evaluate the potential for overfitting in our algorithm.

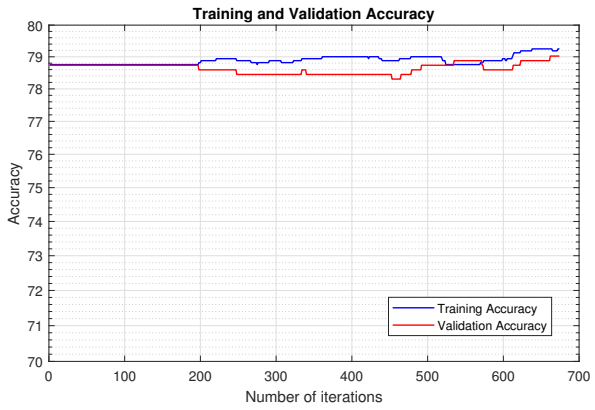

Figure 1: Accuracy plots for the iterations of Algorithm 3.1 (RSLP-L_1).

Figure 2: Loss plots for the iterations of Algorithm 3.1 (RSLP-L_1).

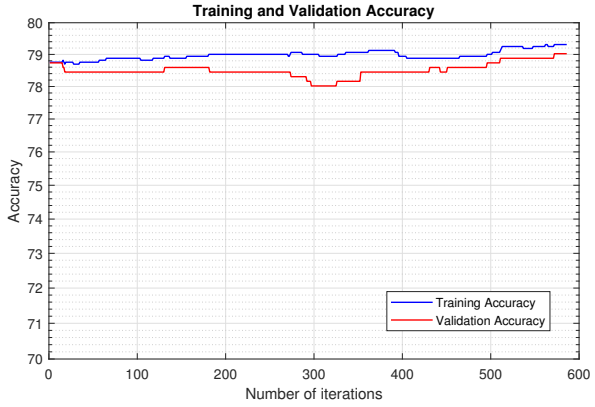

Figure 3: Accuracy plots for the iterations of Algorithm 3.1 (RSLP-L_2).

Figure 4: Loss plots for the iterations of Algorithm 3.1 (RSLP-L_2).

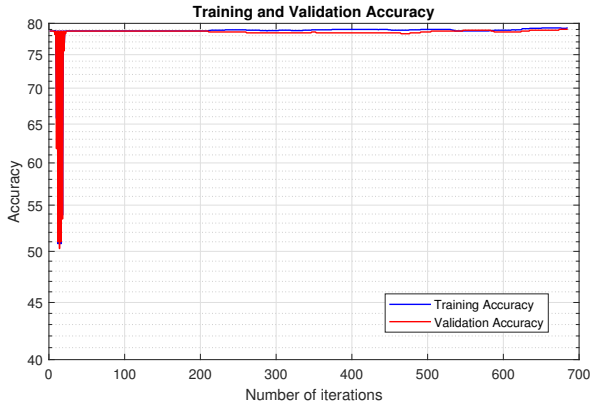

Figure 5: Accuracy plots for the iterations of Algorithm 3.1 (RSLPC-L_1).

Figure 6: Loss plots for the iterations of Algorithm 3.1 (RSLPC-L_1).

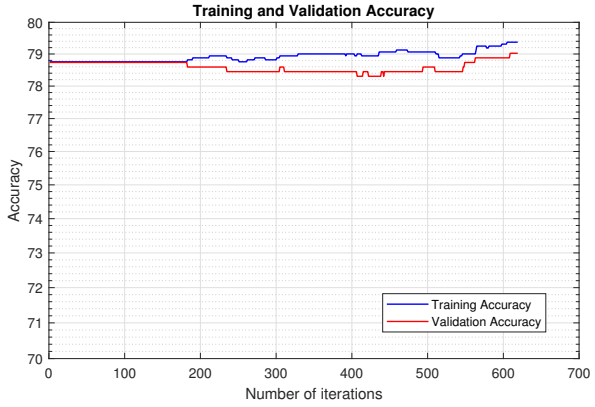

Figure 7: Accuracy plots for the iterations of Algorithm 3.1 (RSLPC-L_2).

Figure 8: Loss plots for the iterations of Algorithm 3.1 (RSLPC-L_2).

Based on Figures 1-8, we can see that the training loss and validation loss values initially decrease until reaching a certain point, after which they stabilize. Conversely, when we analyze the accuracy graph, it becomes evident that both training and validation accuracy show an upward trend, with validation accuracy consistently surpassing training accuracy.

5. Conclusions

This paper introduces a new two inertial projective forward-backward splitting algorithm for solving variational inclusion problems in real Hilbert spaces. We establish a proof of weak convergence for this method, subject to certain mild conditions. Furthermore, we leverage our algorithm as a machine learning technique by incorporating the extreme learning machine model (ELM) to address classification problems. The study outcomes demonstrate that our algorithm outperforms the machine learning methods outlined in Table 5 in terms of efficiency. About the accuracy and loss graphs, it becomes evident from Figures 1-8 that our algorithm does not display any signs of overfitting.

Data availability

The dataset used in this research is publicly available at here.

Institutional review board statement

This study was conducted in accordance with the Declaration of Helsinki, the Belmont report, CIOMS Guideline international conference on Harmonization in Good Cinical Pratice or ICH-GCP or 45CFR 46.101(b) and with approval from the Ethics Committee and Institutional Review Board of Faculty of University of Phayao (Institutional Review Board (IRB) approval, IRB Number: HREC-UP-HSST 1.104066).

Conflicts of interest

The authors declare that they have no conflicts of interest regarding the publication.

Author contributions

Writing this paper and software, C.M.; review and software, P.P.; review medical information, W.L.; review and editing, W.C. All authors have read and agreed to the published version of the manuscript.

Acknowledgment

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

References

[1] M. Akil, R. Saouli, R. Kachouri, *Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy*, Med. Image Anal., **63** (2020). 4
[2] M. Al-Qurashi, S. Rashid, F. Jarad, E. Ali, R. H. Egami, *Dynamic prediction modelling and equilibrium stability of a fractional discrete biophysical neuron model*, Results Phys., **48** (2023), 11 pages. 1
[3] Q. H. Ansari, I. V. Konnov, J. C. Yao, *Existence of a solution and variational principles for vector equilibrium problems*, J. Optim. Theory Appl., **110** (2001), 481–492. 1
[4] M. Bianchi, N. Hadjisavvas, S. Schaible, *Vector equilibrium problems with generalized monotone bifunctions*, J. Optim. Theory Appl., **92** (1997), 527–542. 1
[5] P. L. Combettes, V. R. Wajs, *Signal recovery by proximal forward-backward splitting*, Multiscale Model. Simul., **4** (2005), 1168–1200. 1
[6] J.-Y. Fu, *Generalized vector quasi-equilibrium problems*, Math. Methods Oper. Res., **52** (2000), 57–64. 1
[7] K. Goebel, W. A. Kirk, *Topics in metric fixed point theory*, Cambridge University Press, Cambridge, (1990). 2.2
[8] L. He, *Bone mineral density*, Harvard Dataverse, (2022). 4
[9] O. S. Iyiola, Y. Shehu, *Convergence results of two-step inertial proximal point algorithm*, Appl. Numer. Math., **182** (2022), 57–75. 1
[10] E. N. Khobotov, *Modification of the extra-gradient method for solving variational inequalities and certain optimization problems*, USSR Comput. Math. Math. Phys., **27** (1987), 120–127. 1
[11] P.-L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., **16** (1979), 964–979. 1
[12] G. López, V. Martín-Márquez, F. Wang, H.-K. Xu, *Forward-backward splitting methods for accretive operators in Banach spaces*, Abstr. Appl. Anal., **2012** (2012), 25 pages. 2.3
[13] D. A. Lorenz, T. Pock, *An inertial forward-backward algorithm for monotone inclusions*, J. Math. Imaging Vis., **51** (2015), 311–325. 1
[14] D. T. Luc, N. X. Tan, *Existence conditions in variational inclusions with constraints*, Optimization, **53** (2004), 505–515. 1
[15] A. Moudafi, M. Oliny, *Convergence of a splitting inertial proximal method for monotone operators*, J. Comput. Appl. Math., **155** (2003), 447–454. 1
[16] E. U. Ofoedu, *Strong convergence theorem for uniformly* L*-Lipschitzian asymptotically pseudocontractive mapping in real Banach space*, J. Math. Anal. Appl., **321** (2006), 722–728. 2.5
[17] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc., **73** (1967), 591–597. 2.4
[18] T. Seangwattana, K. Sombut, A. Arunchai, K. Sitthithakerngkiet, *A modified Tseng's method for solving the modified variational inclusion problems and its applications*, Symmetry, **13** (2021), 17 pages. 1
[19] N. X. Tan, *On the existence of solutions of quasivariational inclusion problems*, J. Optim. Theory Appl., **123** (2004), 619–638. 1
[20] L. A. Tuan, P. H. Sach, *Existence of solutions of generalized quasivariational inequalities with set-valued maps*, Acta Math. Vietnam., **29** (2004), 309–316. 1
[21] N. W. S. Wardhani, M. Y. Rochayani, A. Iriany, A. D. Sulistyono, P. Lestantyo, *Cross-validation metrics for evaluating classification performance on imbalanced data*, In: 2019 international conference on computer, control, informatics and its applications (IC3INA), IEEE, (2019), 14–18. 4