

# AUGMENTED DECISION-MAKING IN VENTURE CAPITAL: INTEGRATING DIGITAL TWINS AND AI-BASED MARKET INTELLIGENCE

Zornitsa Yordanova<sup>1</sup>, Hamed Nozari<sup>2</sup>

Received 26.02.2026.

| Sent to review 19.03.2026. | Accepted 13.08.2026.

Original article



<sup>1</sup> University of National and World Economy, Sofia, Industrial Business Department

<sup>2</sup> University of National and World Economy, Sofia, Institute for Economics and Politics, Bulgaria

\*Corresponding Author:  
Zornitsa Yordanova

Email:  
[zornitsayordanova@unwe.bg](mailto:zornitsayordanova@unwe.bg)

JEL Classification:  
G24, O32, D81, C61, C63

Doi: 10.2478/eoik-2026-0047

UDK: 658.14/.17:330.101.541

## ABSTRACT

Venture capital decision-making is increasingly influenced by uncertainty, information asymmetry, and the growing volume of unstructured data. Although previous research has advanced analytical methods for portfolio selection, existing approaches rarely integrate dynamic system behaviour with qualitative market intelligence within a unified framework. This study proposes an AI-augmented decision framework that combines digital twins with large language model-based market intelligence to support venture capital investment under uncertainty. The framework integrates structured financial data with real-time qualitative insights from market narratives, while digital twins simulate startup development and scenario-based performance. Embedded within a multi-criteria decision structure, these inputs enable more adaptive, transparent, and interpretable portfolio selection. Empirical findings indicate that the proposed framework improves decision quality by increasing expected returns while reducing risk and prediction error. It also strengthens decision stability and supports more effective evaluation of trade-offs in complex investment environments. The study contributes to the literature on AI-enabled decision-making and innovation finance by demonstrating how emerging technologies can improve investment evaluation and reduce uncertainty in entrepreneurial ecosystems.

**Keywords:** *Venture Capital, Digital Twin, Large Language Models, Fuzzy Multi-Objective Optimization, Intelligent Decision Making*

## 1. INTRODUCTION

Venture capital (VC) decision-making plays a central role in innovation ecosystems by determining which ventures receive the resources required for growth and technological development. However, these decisions are inherently affected by uncertainty, information asymmetry, and subjective judgement, particularly in early-stage investments where reliable financial data are scarce (Zhang et al., 2026). Investors must assess not only financial indicators but also incomplete and rapidly changing information concerning market conditions, technological potential, and founding-team capabilities (Younus et al., 2026; Nozari et al., 2025a). VC decision-making therefore remains a complex process that combines analytical assessment with contextual and interpretive judgement.

Despite advances in analytical methods, existing approaches to VC evaluation remain limited in their ability to support adaptive and data-rich decisions. Previous research has applied multi-objective optimisation and data-driven models to balance expected returns, risk, and other investment criteria (Bousia, 2026; Kellekci et al., 2025). Although these methods provide valuable analytical structure, they largely depend on structured quantitative data and often neglect the qualitative and contextual information that shapes investment decisions (Dinçer et al., 2025). Moreover, many models adopt a static perspective and therefore fail to capture the dynamic development of startups and their interaction with changing market environments (Wei, 2025).

The increasing availability of unstructured data, including market reports, news, and startup narratives, creates new opportunities to enhance investment evaluation. Recent developments in large language models (LLMs) have improved the extraction of meaningful insights from such data, including signals of market sentiment, perceived risk, and emerging opportunities (Yordanova & Nozari, 2026a; Muneeb et al., 2025). Nevertheless, incorporating these insights into formal decision frameworks remains challenging because of their contextual, ambiguous, and difficult-to-quantify nature (Molavi et al., 2025). Although the value of unstructured data is increasingly recognised, its systematic integration into VC decision-making remains underdeveloped.

In parallel, digital twins have emerged as powerful tools for modelling complex and dynamic systems. By creating virtual representations of real-world entities, they enable scenario analysis, prediction of future behaviour, and continuous updating of system states (Amoozad Mahdiraji et al., 2025; Nozari & Yordanova, 2026b). These capabilities are particularly relevant to VC investment, where outcomes depend on uncertain and evolving technological, organisational, and market conditions (Meharroof et al., 2025). However, the application of digital twins in venture capital remains limited, and their integration with data-driven and multi-criteria decision-making approaches has not been sufficiently examined (Gupta et al., 2025).

Taken together, these developments reveal an important gap in the literature. Although previous studies have advanced multi-objective optimisation, unstructured-data analysis, and dynamic system modelling, these research streams have largely developed in isolation. Consequently, there is a lack of integrative frameworks explaining how diverse forms of data and analytical capabilities can be combined to support investment decisions in uncertain and rapidly changing innovation environments. To address this gap, this study develops an AI-augmented decision-making framework that integrates digital twins and LLM-based market intelligence within a unified structure for VC investment evaluation. Qualitative insights extracted from unstructured data are transformed into interpretable decision inputs, while digital twins provide dynamic representations of startup development and enable scenario-based assessment. By incorporating these components into a multi-criteria decision environment, the framework supports more adaptive, transparent, and informed portfolio choices.

This study makes three principal contributions. First, it advances research on AI-enabled decision-making in innovation finance by demonstrating how unstructured data can be systematically incorporated into investment evaluation. Second, it introduces digital twins as a dynamic decision-support mechanism for venture capital, extending their application beyond traditional engineering and operational contexts. Third, it offers an integrated approach that combines analytical rigour with contextual interpretation, thereby addressing the limitations of exclusively quantitative or qualitative methods. From a managerial perspective, the proposed framework helps investors navigate the complexity of VC decisions by integrating diverse data sources and supporting scenario-based reasoning. It facilitates a shift from static, experience-driven judgement towards more adaptive, data-informed, and explainable decision processes.

The remainder of the paper is structured as follows. The next section reviews the literature on VC decision-making, AI-based analytics, and dynamic modelling. The subsequent sections

present the proposed framework, methodology, empirical analysis, and results. The final section discusses the study's implications, limitations, and directions for future research.

## 2. LITERATURE REVIEW

Recent venture capital (VC) research has increasingly adopted advanced quantitative and data-driven approaches (Chattham et al., 2025). Multi-criteria decision-making and multi-objective optimisation have become important tools for addressing trade-offs among return, risk, liquidity, and other investment objectives (Rayhan et al., 2025; Demir & Chatterjee, 2026). Evolutionary algorithms, in particular, enable the generation of Pareto-optimal solutions from which investors can select according to their preferences (Ibazebo et al., 2025). However, most existing models rely primarily on structured quantitative data and give limited consideration to qualitative and unstructured information (Aydognmus & Aydogmus, 2025).

Fuzzy logic has been introduced to represent uncertainty and ambiguity in decision criteria (Nozari & Yordanova, 2025). By expressing subjective judgements and incomplete information through linguistic variables and membership functions, fuzzy approaches can better reflect real-world decision conditions (Afolabi et al., 2025). This is particularly relevant to VC evaluation because criteria such as management-team quality, innovativeness, and market attractiveness are inherently qualitative (Yin et al., 2025). Nevertheless, most fuzzy models remain static and cannot adequately capture temporal variation or environmental dynamics (Papadopoulou et al., 2026).

Digital twins offer a potential solution to this limitation by creating virtual representations of real-world systems that support scenario analysis, behavioural prediction, and assessment of alternative decisions (Rahim et al., 2025). Although digital twins have been widely applied in industrial and supply-chain settings to improve efficiency and reduce uncertainty (Wang et al., 2025), their use in VC decision-making remains limited and largely conceptual. In particular, operational frameworks integrating digital twins with multi-objective decision models are still underdeveloped (Nozari, 2025).

Advances in natural language processing and large language models (LLMs) have also expanded the capacity to analyse unstructured information (Hirankittiwong et al., 2025). LLMs can process large volumes of news, reports, social media content, and other textual data to identify market sentiment, emerging trends, and risk signals (Castronovo et al., 2025). Their application in finance and investment has therefore created new opportunities for incorporating qualitative market intelligence into decision-making (Gupta et al., 2026). A major challenge, however, is converting contextual and uncertain textual outputs into variables suitable for quantitative modelling (Chen, 2026). Much of the existing research remains focused on textual analysis, while the operational integration of LLM-generated insights into optimisation frameworks is still limited (Ayçin & Erarslan, 2025).

Some studies have sought to address these limitations by combining machine-learning and optimisation techniques within hybrid decision frameworks. These approaches generally aim to improve predictive accuracy and solution quality (Nozari et al., 2025b). Nevertheless, their components often operate independently, with limited interaction among prediction, evaluation, and optimisation layers. Many models also face interpretability challenges and do not clearly explain how particular data sources and criteria influence final investment decisions (Li et al., 2025).

Overall, substantial progress has been achieved in multi-objective optimisation, fuzzy modelling, digital twins, and LLM-based analysis, but these research streams have largely developed in isolation. Existing models rarely analyse structured and unstructured information simultaneously or incorporate contextual intelligence into a dynamic and interpretable VC decision framework.

Three principal research gaps therefore remain. First, there is no well-developed operational mechanism for incorporating LLM-generated market intelligence into multi-objective optimisation models. Second, digital twins have not been effectively employed to support dynamic parameter updating and scenario-based evaluation in VC investment. Third, existing frameworks do not adequately integrate structured quantitative data and unstructured qualitative information within a unified decision process. To address these gaps, this study proposes an integrated framework in which LLM-derived information is represented through fuzzy variables, while a digital twin provides a dynamic layer for parameter updating and scenario analysis. These components are embedded within a multi-objective decision structure, enabling more accurate, adaptive, transparent, and interpretable investment evaluation. The proposed framework thus contributes to the development of intelligent decision-support systems for venture capital under uncertainty.

### 3. CONCEPTUAL FRAMEWORK

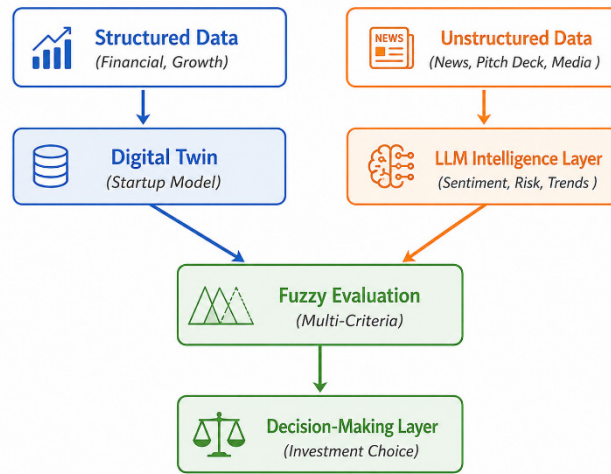
This study proposes an integrated conceptual framework for venture capital decision support that combines three core components: digital twin modelling, large language model (LLM)-based analysis of unstructured data, and fuzzy multi-criteria evaluation. The framework addresses the limitations of conventional decision approaches by enabling the simultaneous analysis of quantitative and qualitative information within a dynamic decision environment. The proposed framework follows a multi-layer architecture in which data from diverse sources are collected, processed, integrated, and transformed into investment recommendations. Each startup is treated as a dynamic entity whose development can be modelled, predicted, and evaluated over time. The input layer incorporates structured data, including financial indicators, growth metrics, and operational information, as well as unstructured data derived from news, market reports, social media, pitch decks, and other startup-related narratives. These data are processed within the analytical core of the framework, which consists of two principal components: a digital twin and an LLM-based market intelligence engine. Their outputs are subsequently integrated within a fuzzy evaluation layer and incorporated into a multi-criteria decision-making mechanism. This structure enables investors to identify preferred investment opportunities while accounting for multiple and potentially conflicting objectives.

A digital twin is constructed for each startup as a dynamic, multidimensional representation of its current and expected performance. It is based on variables covering several dimensions, including financial performance, market development, management-team characteristics, and technological maturity. Relevant indicators include cash burn rate, financial runway, user growth, market share, team experience and stability, and technology-readiness level.

The principal function of the digital twin is to simulate future startup behaviour under alternative conditions. It can therefore assess scenarios such as changes in market demand, additional financing, shifts in operating costs, or variations in growth trajectories. By evaluating the potential effects of these scenarios on startup performance, the digital twin extends investment assessment beyond static analysis towards predictive and adaptive decision support. The LLM-based component functions as a cognitive market intelligence layer that processes extensive volumes of unstructured textual information and converts them into decision-relevant indicators. By analysing news, social media content, industry reports, and pitch decks, it identifies market sentiment, emerging trends, risk signals, and the quality and credibility of startup narratives. These outputs are transformed into quantitative or semi-quantitative variables that can be assessed alongside structured financial and operational data. This component reduces information asymmetry and provides investors with a more comprehensive understanding of both startup conditions and external market perceptions. Given the ambiguity and uncertain-

ty inherent in startup evaluation, deterministic models alone are insufficient to capture the complexity of VC decision-making. The framework therefore applies fuzzy logic to represent imprecise assessments. Key criteria, including expected return, risk, scalability, technological potential, and team quality, are expressed through linguistic variables such as “low,” “medium,” and “high.” Appropriate membership functions then convert these qualitative assessments into values suitable for analytical processing. The fuzzy evaluation layer integrates outputs from the digital twin and the LLM-based intelligence engine, enabling a comprehensive and flexible assessment of each investment opportunity. The resulting evaluations provide the basis for the subsequent multi-criteria decision process and support more adaptive, transparent, and context-sensitive venture capital investment decisions.

Figure 1. Conceptual architecture of a venture capital decision-making system



Source: Authors

Figure 1 illustrates the overall information flow within the proposed framework. Structured and unstructured data enter the system through parallel processing pathways. Structured financial, operational, and market data are incorporated directly into the digital twin, whereas unstructured textual data are processed by large language models and transformed into decision-relevant indicators.

The outputs of these two pathways are subsequently integrated within the fuzzy evaluation layer, which combines quantitative evidence with qualitative and subjective assessments under conditions of uncertainty. The resulting evaluations are then transferred to the decision-making layer, where they support the comparative assessment and optimal selection of venture capital investment opportunities.

#### 4. MATHEMATICAL MODELING

This study develops an integrated mathematical model for venture capital decision-making that optimally selects startups and allocates capital under conditions of uncertainty, multiple objectives, and heterogeneous data. Within the proposed framework, structured financial and operational data are represented through digital twins, while unstructured information from market reports, news, pitch decks, and broader environmental signals is processed by large language models and converted into decision-relevant indicators.

Venture capital decisions are not determined solely by expected financial returns. They also depend on factors such as investment risk, management-team quality, technological maturity, market attractiveness, exit potential, strategic fit, and portfolio diversification. To capture

these interrelated considerations, the proposed model is formulated as a fuzzy mixed-integer multi-objective programming framework that incorporates both continuous and binary decision variables. This formulation enables a more realistic representation of the complexity, ambiguity, and trade-offs inherent in venture capital investment decisions.

### Sets:

- $i \in I$ : Set of candidate startups.
- $s \in S$ : Set of industry sectors.
- $g \in G$ : Set of investment stages (e.g., Seed, Early, Growth).
- $t \in T$ : Set of time periods.
- $\omega \in \Omega$ : Set of market/environmental scenarios.
- $k \in K$ : Set of evaluation criteria.
- $I_s \subseteq I$ : Subset of startups belonging to sector  $s$ .
- $I_g \subseteq I$ : Subset of startups belonging to stage  $g$ .
- $\bar{P} \subseteq I \times I$ : Set of incompatible or mutually exclusive startup pairs.

### Parameters:

- $B$ : Total available investment budget.
- $B_t$ : Budget available in period  $t$ .
- $L_i$ : Minimum investment level for startup  $i$ .
- $U_i$ : Maximum investment level for startup  $i$ .
- $C_i$ : Fixed evaluation/transaction cost for startup  $i$ .
- $\delta_t$ : Discount factor for period  $t$ .
- $p_\omega$ : Probability of scenario  $\omega$ .
- $R_{i\omega}$ : Expected return of startup  $i$  in period  $t$  under scenario  $\omega$  (estimated via Digital Twin).
- $D_{i\omega}$ : Downside loss indicator for startup  $i$  under scenario  $\omega$ .
- $\Gamma_{i\omega}$ : Resilience index of startup  $i$  under scenario  $\omega$ .
- $\sigma_i$ : Risk score of startup  $i$ .
- $M_i$ : Market intelligence score derived from LLM for startup  $i$ .
- $SF_i$ : Strategic alignment score of startup  $i$ .
- $TQ_i$ : Team quality score of startup  $i$ .
- $TR_i$ : Technology readiness score of startup  $i$ .
- $EX_i$ : Exit potential score of startup  $i$ .
- $\alpha_i \in \{0,1\}$ : Eligibility indicator; equals 1 if startup  $i$  is eligible, otherwise 0.
- $B_s$ : Maximum allowable budget for sector  $s$ .
- $N_s$ : Maximum number of startups in sector  $s$ .
- $\bar{B}_g$ : Maximum allowable budget for stage  $g$ .
- $N_g$ : Maximum number of startups in stage  $g$ .
- $N^{min}$ : Minimum portfolio size.
- $N^{max}$ : Maximum portfolio size.
- $S^{min}$ : Minimum number of active sectors.
- $G^{min}$ : Minimum number of active stages.
- $\phi_i$ : Maximum allowable budget share for startup  $i$ .
- $Risk^{max}$ : Maximum acceptable portfolio risk.
- $Loss_\omega^{max}$ : Maximum allowable downside loss under scenario  $\omega$ .

$MI^{min}$ : Minimum acceptable portfolio-level market intelligence score.

$SF^{min}$ : Minimum acceptable strategic alignment threshold.

$TQ^{min}$ : Minimum acceptable team quality threshold.

$TR^{min}$ : Minimum acceptable technology readiness threshold.

$EX^{min}$ : Minimum acceptable exit potential threshold.

$w_k$ : Weight of criterion  $k$ .

$Z_j^{best}$ : Best value of objective  $j$ .

$Z_j^{worst}$ : Worst value of objective  $j$ .

### Decision Variables:

$x_{it} \geq 0$ : Investment allocated to startup  $i$  in period  $t$ .

$y_i \in \{0,1\}$ : Equals 1 if startup  $i$  is selected; otherwise 0.

$u_s \in \{0,1\}$ : Equals 1 if sector  $s$  is active; otherwise 0.

$v_g \in \{0,1\}$ : Equals 1 if stage  $g$  is active; otherwise 0.

$\lambda \in [0,1]$ : Overall satisfaction level.

$\mu_j \in [0,1]$ : Satisfaction degree of objective  $j$ .

### Objective Functions:

$$\max Z_1 = \sum_i \sum_t \sum_\omega \delta_i p_{\omega} R_{it\omega} x_{it} - \sum_i C_i y_i \quad (1)$$

$$\min Z_2 = \sum_i \sigma_i \left( \sum_t x_{it} \right) \quad (2)$$

$$\max Z_3 = \sum_i (\beta_1 MI_i + \beta_2 SF_i + \beta_3 TQ_i) \left( \sum_t x_{it} \right) \quad (3)$$

$$\max Z_4 = \sum_i \sum_t \sum_\omega p_{\omega} \Gamma_{i\omega} x_{it} \quad (4)$$

$$\min Z_5 = \sum_i \sum_\omega p_{\omega} D_{i\omega} \left( \sum_t x_{it} \right) \quad (5)$$

### Fuzzy Aggregation of Objectives:

Since the above objective functions may naturally conflict with each other, a fuzzy satisfaction approach is used to integrate them. For each objective function, a membership function is defined.

For maximization objectives:

$$\mu_j(Z_j) = \frac{Z_j - Z_j^{worst}}{Z_j^{best} - Z_j^{worst}} \quad \forall j \in J^{\max} \quad (6)$$

For minimization objectives:

$$\mu_j(Z_j) = \frac{Z_j^{worst} - Z_j}{Z_j^{worst} - Z_j^{best}} \quad \forall j \in J^{\min} \quad (7)$$

The final fuzzy model is then formulated as follows:

$$\begin{aligned} & \max \lambda \\ & \text{subject to} \\ & \lambda \leq \mu_j(Z_j) \quad \forall j \end{aligned} \quad (8)$$

**S.t:**

$$\sum_{i \in I} \sum_{t \in T} x_{it} + \sum_{i \in I} c_i y_i \leq B \quad (9)$$

$$\sum_{i \in I} x_{it} \leq B_t \quad \forall t \in T \quad (10)$$

$$\sum_{t \in T} x_{it} \geq L y_i \quad \forall i \in I \quad (11)$$

$$\sum_{t \in T} x_{it} \leq U y_i \quad \forall i \in I \quad (12)$$

$$\sum_{i \in I} y_i \geq N^{\min} \quad (13)$$

$$\sum_{i \in I} y_i \geq N^{\max} \quad (14)$$

$$\sum_{t \in T} x_{it} \leq \phi_i B \quad \forall i \in I \quad (15)$$

$$\sum_{i \in I_s} \sum_{t \in T} x_{it} \leq \bar{B}_s \quad \forall s \in S \quad (16)$$

$$\sum_{i \in I_s} y_i \leq \bar{N}_s \quad \forall s \in S \quad (17)$$

$$\sum_{i \in I_g} \sum_{t \in T} x_{it} \leq \bar{B}_g \quad \forall g \in G \quad (18)$$

$$\sum_{i \in I_g} y_i \leq \bar{N}_g \quad \forall g \in G \quad (19)$$

$$y_i \leq u_s \quad \forall s \in S, \quad \forall i \in I_s \quad (20)$$

$$\sum_{i \in I_s} y_i \geq u_s \quad \forall s \in S \quad (21)$$

$$\sum_{s \in S} u_s \geq S^{\min} \quad (22)$$

$$\sum_{i \in I_g} y_i \leq I_g \mid v_g \quad \forall g \in G \quad (23)$$

$$\sum_{i \in I_g} y_i \geq v_g \quad \forall g \in G \quad (24)$$

$$\sum_{g \in G} v_g \geq G^{\min} \quad (25)$$

$$y_i \leq \alpha_i \quad \forall i \in I \quad (26)$$

$$\sum_{i \in I} \sigma_i \left( \sum_{t \in T} x_{it} \right) \leq Risk^{\max} \quad (27)$$

$$\sum_{i \in I} D_{i\omega} \left( \sum_{t \in T} x_{it} \right) \leq Loss_w^{\max} \quad \forall \omega \in \Omega \quad (28)$$

$$\sum_{i \in I} M I_i \left( \sum_{t \in T} x_{it} \right) \geq M I^{\min} \sum_{i \in I} \sum_{t \in T} x_{it} \quad (29)$$

$$\sum_{i \in I} S F_i \left( \sum_{t \in T} x_{it} \right) \geq S F^{\min} \sum_{i \in I} \sum_{t \in T} x_{it} \quad (30)$$

$$\sum_{i \in I} TQ_i \left( \sum_{t \in T} x_{it} \right) \geq TQ^{\min} \sum_{i \in I} \sum_{t \in T} x_{it} \quad (31)$$

$$\sum_{i \in I} TR_i \left( \sum_{t \in T} x_{it} \right) \geq TR^{\min} \sum_{i \in I} \sum_{t \in T} x_{it} \quad (32)$$

$$\sum_{i \in I} EX_i \left( \sum_{t \in T} x_{it} \right) \geq EX^{\min} \sum_{i \in I} \sum_{t \in T} x_{it} \quad (33)$$

### Variable Domains:

$$x_{it} \geq 0 \quad \forall i \in I, \forall t \in T \quad (34)$$

$$y_i \in \{0,1\} \quad \forall i \in I$$

$$u_s \in \{0,1\} \quad \forall s \in S$$

$$v_g \in \{0,1\} \quad \forall g \in G$$

$$0 \leq \lambda \leq 1$$

$$0 \leq \mu_j \leq 1 \quad \forall j$$

Objective function (1) maximizes the discounted expected return of the investment portfolio. It seeks to generate the highest possible economic benefit by considering startup-specific returns across time periods and scenarios. Fixed costs associated with entry, evaluation, and transaction execution are deducted to provide a more realistic estimate of net investment returns.

Objective function (2) minimizes total portfolio risk. It determines the selection of startups and the allocation of capital so that aggregate portfolio risk remains as low as possible. This objective discourages excessive exposure to high-risk ventures and supports an appropriate balance between expected return and portfolio stability.

Objective function (3) maximizes market intelligence and strategic alignment. It prioritizes startups that exhibit favourable market signals, strong external perceptions, strategic compatibility with the investor's objectives, and high management-team quality. Accordingly, the function extends the evaluation beyond financial attractiveness by incorporating qualitative and strategic dimensions.

Objective function (4) maximizes portfolio resilience based on digital twin outputs. It favours startups that demonstrate greater adaptability and sustainability under market fluctuations, environmental changes, and adverse scenarios. The function therefore broadens the investment perspective beyond short-term profitability by incorporating long-term survival, adaptability, and recovery capacity.

Objective function (5) minimizes the downside vulnerability of the portfolio. It controls potential losses under adverse scenarios and allocates investments in a manner that reduces both the magnitude and frequency of negative outcomes. Unlike the general risk objective, this function focuses specifically on severe losses and unfavourable portfolio performance.

Constraint (9) ensures that total investment expenditure, including the fixed costs associated with startup selection, does not exceed the available budget. It establishes the overall financial boundary of the model and ensures that investment decisions remain consistent with the investor's available resources.

Constraint (10) imposes a budget limit for each time period. It ensures that investments made during a particular period do not exceed the capital available at that time, thereby incorporating cash-flow and temporal funding limitations into the model.

Constraint (11) specifies the minimum investment required for each selected startup. If a startup is included in the portfolio, it must receive at least the minimum amount of capital neces-

sary for the investment to be economically meaningful.

Constraint (12) defines the maximum investment permitted for each startup. It prevents excessive allocation to a single venture and reduces undesirable capital concentration.

Constraint (13) establishes the minimum portfolio size by requiring the selection of at least a specified number of startups. Its purpose is to prevent the creation of an excessively small and vulnerable portfolio.

Constraint (14) determines the maximum portfolio size. It ensures that the number of selected startups remains within the investor's financial, managerial, and monitoring capacity.

Constraint (15) limits the proportion of the total budget allocated to any individual startup. This restriction reduces portfolio dependence on a single venture and strengthens diversification and concentration-risk management.

Constraint (16) imposes a maximum budget allocation for each industry sector. It prevents a single sector from dominating the portfolio and limits exposure to sector-specific risks and fluctuations.

Constraint (17) limits the number of selected startups within each sector. It complements Constraint (16) by promoting sectoral diversification in terms of portfolio composition rather than capital allocation alone.

Constraint (18) establishes the maximum amount of capital that may be allocated to each investment stage. It prevents excessive concentration of resources in a particular stage, such as seed, early-stage, or growth-stage investment.

Constraint (19) limits the number of startups selected at each investment stage. It promotes diversification across different levels of business maturity and reduces stage-specific concentration.

Constraint (20) links startup selection to the activation of the corresponding sector. A sector can be designated as active only when at least one startup from that sector is selected.

Constraint (21) ensures that each active sector contains at least one selected startup. Together, Constraints (20) and (21) establish consistency between sector-activation variables and startup-selection decisions.

Constraint (22) requires the portfolio to include a minimum number of active sectors. It therefore guarantees a specified level of sectoral diversification.

Constraint (23) links startup selection to the activation of the corresponding investment stage. An investment stage cannot be classified as active unless at least one startup from that stage is selected.

Constraint (24) ensures that every active investment stage includes at least one selected startup. Together, Constraints (23) and (24) maintain consistency between stage-activation and startup-selection variables.

Constraint (25) requires the portfolio to cover a minimum number of active investment stages. It guarantees diversification across different levels of startup maturity.

Constraint (26) imposes an initial eligibility requirement. Only startups that satisfy the predetermined minimum criteria may enter the selection process, thereby excluding unsuitable alternatives before portfolio optimization.

Constraint (27) limits total portfolio risk to a predefined acceptable threshold. It ensures that the pursuit of higher returns does not result in aggregate risk exceeding the investor's tolerance.

Constraint (28) restricts downside losses under each scenario. It strengthens portfolio resilience to adverse market conditions and external shocks by preventing the selection of portfolios that could experience unacceptable losses.

Constraint (29) establishes the minimum acceptable level of portfolio market intelligence. It requires the weighted average of scores generated through LLM-based analysis to remain

above a specified threshold, ensuring that the selected portfolio is supported by favourable market perceptions, trends, and external signals.

Constraint (30) imposes a minimum level of strategic alignment. It ensures that the selected startups are sufficiently compatible with the investor's strategic priorities, investment orientation, and long-term objectives.

Constraint (31) requires a minimum level of management-team quality across the selected startups. This reflects the importance of managerial competence, experience, and execution capacity as central determinants of venture success.

Constraint (32) establishes the minimum technological readiness of the portfolio. It ensures that selected startups demonstrate an acceptable level of technological maturity, product development, and implementation capability.

Constraint (33) imposes a minimum level of exit potential and portfolio liquidity. It ensures that selected startups possess sufficient prospects for successful and timely exit, thereby supporting both value creation and capital realization.

## 5. LLM-FUZZY INTEGRATION MECHANISM

A fundamental challenge in venture capital decision-making is the integration of structured and unstructured data within a coherent analytical framework. Although quantitative information, such as financial and operational indicators, can be readily incorporated into mathematical models, a substantial proportion of decision-relevant evidence remains textual, qualitative, and contextual. Such evidence includes market perceptions, media coverage, pitch deck content, competitive narratives, and behavioural signals related to founding teams, all of which are often overlooked in conventional evaluation models.

To address this limitation, this study proposes an integrated mechanism that combines large language model (LLM) outputs with fuzzy logic. The mechanism functions as an intermediate analytical layer linking LLM-based interpretation of unstructured data with the quantitative evaluation performed by the mathematical decision model.

In the first stage, textual data are collected from multiple sources, including industry news, market analyses, social media, technical reports, and documents submitted by startups. These data are processed by an LLM to perform semantic analysis, contextual interpretation, and feature extraction.

In the second stage, the LLM generates decision-relevant indicators such as market sentiment, perceived risk, competitive intensity, growth potential, narrative quality, and strategic coherence. Because these outputs are inherently contextual, uncertain, and imprecise, they cannot be incorporated directly into a deterministic decision model.

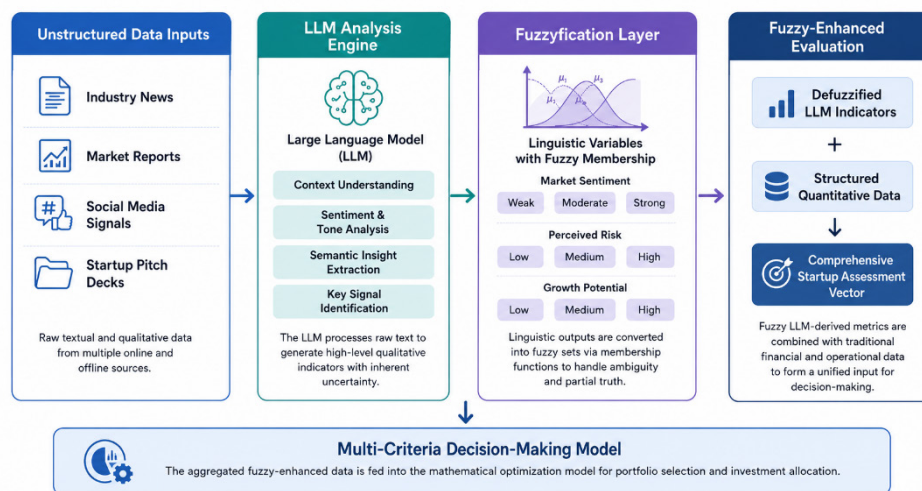
In the third stage, the LLM-generated indicators are transformed into linguistic variables through a structured scoring procedure. Each indicator is first assigned a normalized semantic score based on factors such as sentiment intensity, risk-related language, growth signals, competitive positioning, and the consistency of market narratives. The resulting score is then mapped onto linguistic categories, such as weak, moderate, and strong, using predefined fuzzy membership ranges. Membership functions subsequently convert these linguistic assessments into quantitative values suitable for mathematical processing.

This fuzzification process preserves the uncertainty and ambiguity inherent in textual information while making AI-generated insights analytically usable. The resulting fuzzy values are combined with structured financial, operational, and performance indicators to form a comprehensive evaluation vector for each startup. This vector is then used as input to the multi-crite-

ria decision-making layer.

Accordingly, the final investment decision is based not only on financial and operational performance but also on qualitative market intelligence and contextual analysis. Importantly, LLM outputs are not introduced into the mathematical model without interpretation. Instead, they are systematically structured, normalized, and aligned with the remaining model components through the fuzzification process. This procedure improves the consistency, reliability, transparency, and interpretability of the resulting investment evaluations.

Figure 2. LLM–Fuzzy Integration Mechanism in Venture Capital Decision-Making



Source: Authors

As can be seen in Figure 2, the information flow starts from unstructured data and after passing through the analysis layer based on the linguistic model, it is transformed into a set of semantic indicators. These indicators are then fed into the fuzzification section and represented in the form of linguistic variables and membership functions. The output of this stage is a set of processable values that can be used directly in the decision-making model.

## 6. SOLUTION APPROACHES

The proposed model generates a complex, nonlinear search space because of its multiple objectives, binary and continuous decision variables, fuzzy evaluation functions, and numerous constraints. Consequently, exact optimisation methods become computationally demanding, particularly for medium- and large-scale problem instances. To solve the model, three multi-objective metaheuristic algorithms were employed: Multi-Objective Particle Swarm Optimisation (MOPSO), Multi-Objective Grey Wolf Optimisation (MOGWO), and the Non-Dominated Sorting Genetic Algorithm II (NSGA-II). These algorithms were selected because of their demonstrated ability to balance convergence towards the Pareto front with the preservation of solution diversity in complex multi-objective problems.

Each chromosome, particle, or search agent was represented as a decision vector containing both the binary selection status of each startup and the amount of capital allocated to it across different time periods. For each candidate solution, the objective-function values were calculated using the proposed mathematical model. Constraint violations were managed through a penalty mechanism, whereby infeasible solutions received lower fitness values. This approach directed the search simultaneously towards feasible regions and higher-quality solutions.

MOPSO is based on the collective movement of particles within the search space. Each particle

represents a candidate solution, and its trajectory is influenced by two sources of information: its best previously identified position and a leader selected from an archive of non-dominated solutions. Unlike single-objective particle swarm optimisation, MOPSO maintains a set of Pareto-optimal candidates, enabling the simultaneous exploration of multiple promising regions of the search space (Rivero-Iglesias et al., 2025).

In the present implementation, each particle was assigned an initial position and velocity. An external archive was then used to retain non-dominated solutions, from which leaders were selected to guide particle movement. To prevent excessive concentration in particular regions of the Pareto front, the leader-selection mechanism favoured less densely populated areas and thereby preserved solution diversity. Following each position update, variable bounds were enforced and penalties were applied to constraint violations to maintain feasible and interpretable solutions. The general implementation procedure is presented in Algorithm 1

Algorithm 1. Pseudocode of MOPSO

|                                                                 |
|-----------------------------------------------------------------|
| 1: Initialize particle positions and velocities                 |
| 2: Evaluate all particles based on objective functions          |
| 3: Set personal best for each particle                          |
| 4: Create external archive with non-dominated solutions         |
| 5: while stopping criterion is not satisfied do                 |
| 6:     Select a leader from the archive                         |
| 7:     Update velocity of each particle                         |
| 8:     Update position of each particle                         |
| 9:     Repair infeasible solutions if necessary                 |
| 10:    Apply penalty mechanism for constraint violations        |
| 11:    Re-evaluate particles                                    |
| 12:    Update personal best positions                           |
| 13:    Update external archive with new non-dominated solutions |
| 14: end while                                                   |
| 15: Return the archive as the approximated Pareto front         |

Pseudo-Algorithm 1 illustrates that MOPSO operates through the iterative movement of particles, the retention of each particle's best individual experience, and the use of an external archive to preserve non-dominated solutions. This structure combines individual memory with collective search information. Consequently, rather than producing a single optimal solution, MOPSO generates a diverse set of Pareto-optimal alternatives for multi-objective decision problems.

Multi-Objective Grey Wolf Optimisation (MOGWO) is inspired by the hierarchical structure and hunting behaviour of grey wolves. In the standard algorithm, the three highest-ranked solutions alpha, beta and delta guide the movement of the remaining search agents. Each wolf updates its position according to the locations of these three leaders. In the multi-objective version, leaders are selected from an archive of non-dominated solutions, allowing the search to be guided by several high-quality alternatives rather than a single optimal point (Shang et al., 2025).

In the proposed model, each wolf represents a candidate investment portfolio. Following the generation of the initial population, non-dominated solutions are identified and stored in an external archive, from which the alpha, beta, and delta leaders are selected. The remaining wolves then update their positions relative to these leaders. At each iteration, decision-variable bounds and model constraints are checked, while infeasible solutions are repaired or penalised

to maintain consistency with the mathematical formulation. The implementation procedure is summarised in Algorithm 2.

Algorithm 2. Pseudocode of MOGWO

|                                                                         |
|-------------------------------------------------------------------------|
| 1: Initialize grey wolf population                                      |
| 2: Evaluate all wolves using objective functions                        |
| 3: Identify non-dominated solutions                                     |
| 4: Select alpha, beta, and delta leaders                                |
| 5: while stopping criterion is not satisfied do                         |
| 6:     Update the position of each wolf based on alpha, beta, and delta |
| 7:     Repair infeasible solutions if necessary                         |
| 8:     Apply penalty mechanism for constraint violations                |
| 9:     Re-evaluate wolves                                               |
| 10:    Update the set of non-dominated solutions                        |
| 11:    Re-select alpha, beta, and delta                                 |
| 12: end while                                                           |
| 13: Return the set of non-dominated solutions                           |

Pseudo-Algorithm 2 shows that MOGWO is based on population guidance by three leading solutions and the iterative adjustment of the remaining search agents. This mechanism balances exploration and exploitation by directing the search towards multiple reference solutions rather than a single central optimum, which is particularly suitable for multi-objective problems (Sultan et al., 2025).

The Non-Dominated Sorting Genetic Algorithm II (NSGA-II) is one of the most widely applied methods for solving multi-objective optimisation problems. It relies on two principal mechanisms: non-dominated sorting and crowding distance. Non-dominated sorting classifies candidate solutions into successive Pareto fronts, while crowding distance measures the density of neighbouring solutions and helps preserve diversity. NSGA-II therefore prioritises solutions with superior Pareto ranks while preventing excessive concentration in particular regions of the Pareto front.

In adapting NSGA-II to the proposed model, each individual was represented as a decision vector combining startup-selection variables and corresponding investment allocations. After evaluating the initial population, solutions were ranked through non-dominated sorting, and their crowding distances were calculated. Selection, crossover, and mutation operators were then applied to generate an offspring population.

At each iteration, the parent and offspring populations were combined and re-evaluated. The next generation was selected according to non-dominated rank and crowding distance, ensuring both convergence towards the Pareto front and adequate solution diversity. The overall implementation procedure is presented in Algorithm 3.

Algorithm 3. Pseudocode of NSGA-II

```

1: Initialize population
2: Evaluate objective functions for all individuals
3: while stopping criterion is not satisfied do
4:   Perform non-dominated sorting
5:   Calculate crowding distance
6:   Select parents using selection operator
7:   Generate offspring through crossover and mutation
8:   Repair infeasible solutions if necessary
9:   Apply penalty mechanism for constraint violations
10:  Evaluate offspring population
11:  Combine parent and offspring populations
12:  Perform non-dominated sorting on the combined population
13:  Select the next generation based on rank and crowding distance
14: end while
15: Return the final approximated Pareto front

```

Pseudo-algorithm 3 shows that NSGA-II operates on a systematic evolutionary structure in which the quality and diversity of solutions are controlled simultaneously. The simultaneous use of non-dominated sorting and crowding distance makes this algorithm one of the most suitable options for approximating the Pareto front in complex multi-objective problems.

After describing the algorithms, parameter settings were adjusted so that all three methods run under relatively level and comparable conditions. These settings were selected using preliminary tests and an examination of the convergence behavior of the algorithms. To avoid unnecessary enlargement of the table, only the most important control parameters are reported in Table 1.

Table 1. Main Parameter Settings of the Solution Algorithms

| Algorithm       | Main parameters                                                                                                                |
|-----------------|--------------------------------------------------------------------------------------------------------------------------------|
| MOPSO           | Population size = 100; Max iterations = 200; Archive size = 100;<br>Inertia weight = 0.4–0.9; $c_1=1.5$ ; $c_2=1.5$            |
| MOGWO           | Population size = 100; Max iterations = 200; Archive size = 100;<br>Control parameter $\lambda$ decreased linearly from 2 to 0 |
| NSGA-II         | Population size = 100; Max generations = 200; Crossover rate = 0.80;<br>Mutation rate = 0.20                                   |
| Common settings | Independent runs = 30; Programming language = Python 3.11;<br>Validation software = GAMS 43.4                                  |

Source: Authors

The parameters in Table 1 were chosen to maintain fairness in comparison and to allow for the observation of the actual behavior of each algorithm. The population size was considered equal for all three algorithms to reduce the effect of population size differences on the performance comparison. The number of iterations and generations was also set at a level that gave the search enough time to approach the Pareto front. In the case of MOPSO, the inertia weight was considered as an interval to strengthen the balance between global search and local search. In MOGWO, a linear decrease of the control parameter from 2 to 0 was used to gradually transition from exploration to exploitation. In NSGA-II, the crossover and mutation rates were

also set based on common and stable values in the evolutionary algorithm literature.

The computational implementation of the algorithms was performed in Python 3.11. NumPy 1.26.4 was used for numerical calculations and vector operations, Pandas 2.1.4 was used for data organization and management, and Matplotlib 3.8.2 was used for drawing Pareto fronts and analytical graphs. The programs could be executed in the Jupyter Notebook 7.0 environment as well as PyCharm 2023.3, and the final version of the codes was developed in this environment. In order to reduce the effect of randomness of metaheuristic algorithms, each algorithm was run in 30 independent runs for each problem instance, and all runs were performed with the same parameter structure. The stopping criterion for MOPSO and MOGWO was set to 200 iterations and for NSGA-II to 200 generations. In addition, to equalize the experimental conditions, all algorithms were run with an initial population size of 100.

For validation, the model was also implemented in small and medium dimensions in GAMS 43.4 software. In the exact solution, appropriate solvers of the mixed and nonlinear programming problem were used to allow comparison between the solutions obtained from the metaheuristic methods and the reference outputs. The goal of this step was to ensure that the proposed algorithms were not only capable of producing a non-dominated solution set but also behaved acceptably in terms of compatibility with the real structure of the model. The validation was performed on smaller samples, because in these dimensions it was possible to solve the model exactly and make a direct comparison; while for larger dimensions, the main emphasis was placed on the metaheuristic methods.

## 7. RESULTS AND DISCUSSION

The dataset was designed to reflect the characteristics of the venture capital environment while ensuring analytical control, transparency, and reproducibility. Accordingly, the study combined real and simulated data. The real dataset was compiled from publicly available startup and venture capital sources and included indicators such as user growth, estimated revenue, technological maturity, investment stage, and management-team characteristics.

Unstructured textual data, including market news, industry reports, startup descriptions, and pitch documents, were processed using a large language model. The resulting outputs were transformed into quantitative indicators, including market sentiment, perceived risk, and market attractiveness. Additional variables, such as expected return, resilience, and downside loss, were estimated through the digital twin model to capture dynamic relationships among financial, operational, and market factors.

The real dataset covered 120 startups operating in different technology sectors and at various investment stages. Data were collected from public startup databases, venture capital reports, market news, and pitch-related documents. To complement the empirical observations and test the model under a broader range of conditions, simulated data were generated using statistical distributions calibrated to the characteristics of the real dataset. The simulation process preserved plausible correlations among market growth, risk, expected return, and other key variables.

Three scenarios were considered: optimistic, balanced, and pessimistic. These scenarios were defined through variations in market growth, competitive intensity, risk, expected return, and demand conditions. The optimistic scenario represented favourable market conditions, strong demand, and high growth potential. The pessimistic scenario captured increased volatility, declining demand, intensified competition, and higher risk. The balanced scenario represented normal market conditions and served as the reference case. A probability was assigned to each scenario and incorporated into the multi-objective calculations and portfolio-performance evaluation.

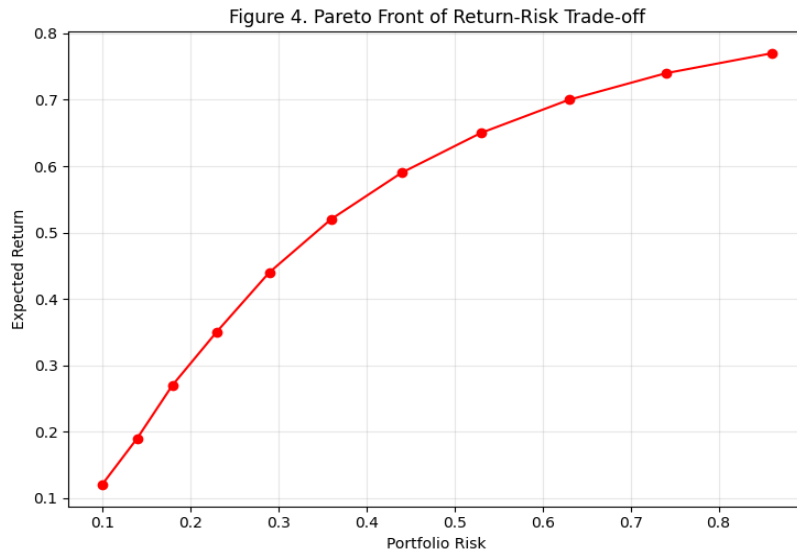
This data structure enabled the model and solution algorithms to be assessed under diverse but realistic conditions. It also facilitated a consistent comparison of algorithmic performance while maintaining transparency regarding the assumptions underlying the empirical analysis.

Pareto-front analysis was conducted to examine the trade-offs among the model's conflicting objec-

tives and evaluate the quality of the generated solutions. Particular attention was given to expected return, portfolio risk, and resilience, as these criteria are central to venture capital investment decisions. The non-dominated solutions generated by the algorithms formed a continuous and interpretable Pareto front representing portfolios with different risk–return–resilience profiles.

Each point on the Pareto front represents a feasible investment portfolio for which improvement in one objective can be achieved only at the expense of at least one other objective. The resulting solution set therefore provides investors with a range of efficient portfolio alternatives rather than prescribing a single universally optimal decision.

Figure 3. Pareto Front of Return–Risk Trade-off in Venture Capital Portfolio

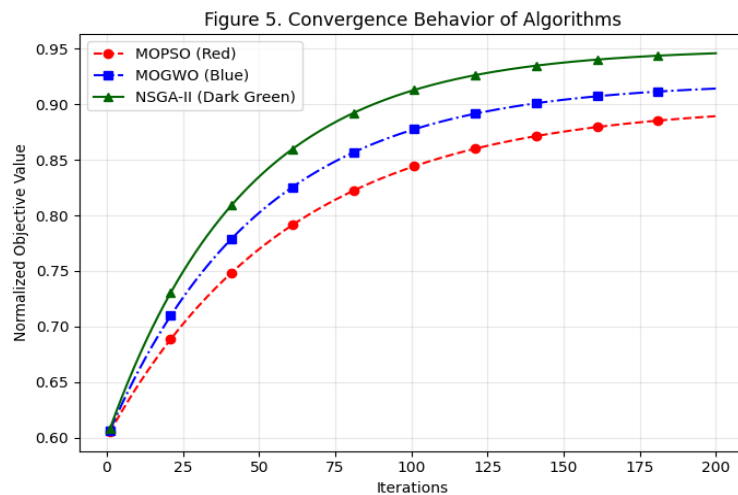


Source: Authors

As can be seen in Figure 3, the Pareto front has an upward trend, indicating that the return increases with increasing risk. In the early regions of the curve, a small increase in risk leads to a significant improvement in returns, while in the later parts, this relationship becomes nonlinear and increasing risk leads to only limited improvements in returns. This behavior indicates the existence of distinct efficient points in the decision space, each of which can be chosen according to the investor’s degree of risk tolerance. In addition, the good dispersion of points on the front indicates that the model is able to generate a diverse set of investment options and that the responses are not concentrated in a particular region.

The performance of the algorithms was compared to evaluate the quality of the generated solutions, their convergence behavior, and their ability to approximate the Pareto front. For this purpose, all three algorithms were run under the same conditions, and the outputs were evaluated based on standard multi-objective criteria including the value of the normalized objective function, the quality of the Pareto front, and computational time. The focus of the analysis was not only on the algorithmic performance, but also on the quality of the solutions within the investment model framework and their ability to produce acceptable portfolios.

Figure 4. Convergence Behavior of Algorithms



Source: Authors

The graph in Figure 4 shows the convergence process of the algorithms over the iterations. As can be seen, all three algorithms start from a similar initial value and gradually move towards more optimal values. The main difference is in the convergence rate and the final value that can be achieved. NSGA-II shows a faster trend in improving the value of the objective function and reaches more suitable regions of the solution space in the initial iterations. MOGWO also has relatively stable behavior and moves towards better solutions with a uniform slope. In contrast, MOPSO, despite good convergence, shows a slower improvement rate than the other two methods. These differences indicate the different ways in which these algorithms search the decision space and their ability to exploit collective information.

Table 2. Comparative Performance of Algorithms

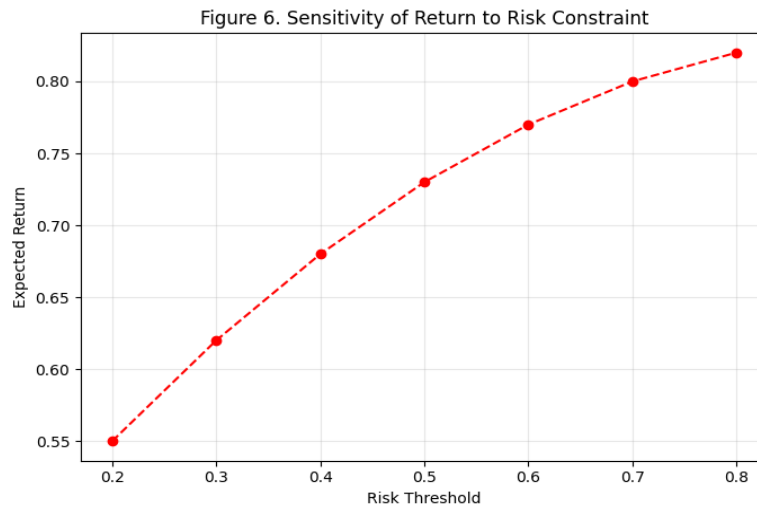
| Algorithm | Best Objective Value | Average Objective Value | Std. Deviation | Computational Time (s) |
|-----------|----------------------|-------------------------|----------------|------------------------|
| MOPSO     | 0.891                | 0.872                   | 0.014          | 12.6                   |
| MOGWO     | 0.915                | 0.901                   | 0.01           | 13.8                   |
| NSGA-II   | 0.946                | 0.932                   | 0.008          | 15.2                   |

Source: Authors

The values presented in Table 2 indicate the relative performance of the algorithms in terms of solution quality and stability. The value of the best objective function indicates the ability of the algorithm to achieve the best points of the Pareto front, while the mean and standard deviation values indicate the stability of performance across different runs. The results show that NSGA-II has the highest objective value and the lowest dispersion, which indicates its higher stability in producing quality solutions. MOGWO also provides balanced performance and is close to NSGA-II in terms of average performance, while MOPSO has more dispersion in results despite shorter execution time.

To assess the stability of the model and examine the impact of changing key parameters on the outputs, a sensitivity analysis was conducted on the most important decision-making parameter, the portfolio risk limit. This parameter plays a direct role in the formation of the investment mix, as it determines the extent to which one can move towards higher yielding but riskier options. In this analysis, the risk limit value was changed over a specified range and its effect on the expected return of the portfolio was examined.

Figure 5. Sensitivity of Return to Risk Constraint



Source: Authors

As can be seen in Figure 5, a gradual increase in the risk limit leads to an increase in the expected return, but this relationship is not completely linear. At low risk levels, a limited increase in the risk limit leads to a significant improvement in return, while at higher levels, the slope of the curve decreases and further increases in risk lead to only minor improvements in return. This behavior indicates the existence of an optimal decision-making zone where a good balance between risk and return is achieved and exceeding it does not significantly increase the efficiency of the investment.

Table 3. Sensitivity Results for Risk Threshold

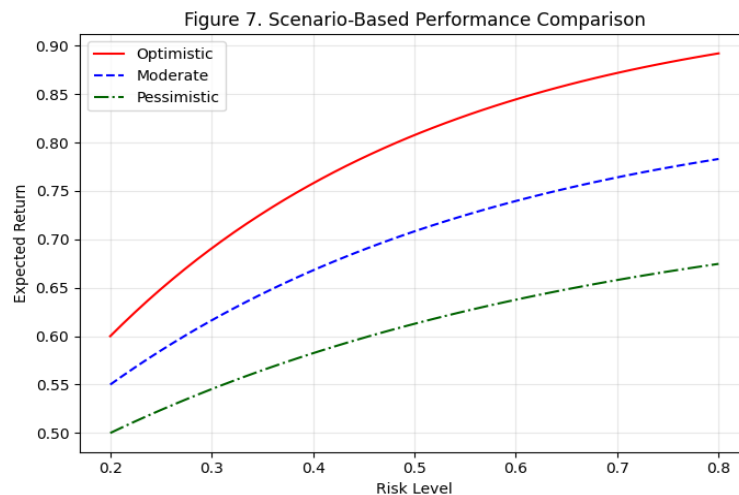
| Risk Threshold | Expected Return | Portfolio Risk | Resilience Index |
|----------------|-----------------|----------------|------------------|
| 0.2            | 0.55            | 0.21           | 0.78             |
| 0.3            | 0.62            | 0.29           | 0.75             |
| 0.4            | 0.68            | 0.38           | 0.72             |
| 0.5            | 0.73            | 0.47           | 0.69             |
| 0.6            | 0.77            | 0.56           | 0.66             |
| 0.7            | 0.80            | 0.64           | 0.63             |
| 0.8            | 0.82            | 0.73           | 0.60             |

Source: Authors

The values in Table 3 quantitatively confirm the trend observed in the graph. As the risk threshold increases, the expected return increases uniformly, while the resilience index decreases. This indicates that choosing portfolios with higher returns is usually associated with reduced resilience to fluctuations. Therefore, the decision maker must choose a point at which the increase in return is still worth accepting the additional risk.

To examine the behavior of the model in different environmental conditions and assess the sustainability of decisions, scenario-based analysis was conducted based on three distinct situations including optimistic, balanced, and pessimistic scenarios (Borisov & Hristozov, 2024). In this analysis, key parameters such as market growth rate, systematic risk level, and expected return were considered differently in each scenario to determine the impact of changing external conditions on portfolio performance. The purpose of this analysis was to measure the flexibility of the model and its ability to produce decisions consistent with changing environmental conditions.

Figure 6. Scenario-Based Performance Comparison



Source: Authors

As can be seen in Figure 6, in the optimistic scenario, the expected return is higher at all risk levels than the other two scenarios, and the slope of the curve also indicates greater growth opportunities in these conditions. In the balanced scenario, the system's behavior is in the middle, and the trend of increasing returns follows a more uniform slope. In contrast, the pessimistic scenario is associated with the lowest level of return, and increasing risk in this case leads to a more limited improvement in portfolio performance. The distance between the curves indicates the sensitivity of the model to market conditions and the importance of considering uncertainty in decision-making.

Table 4. Scenario-Based Performance Metrics

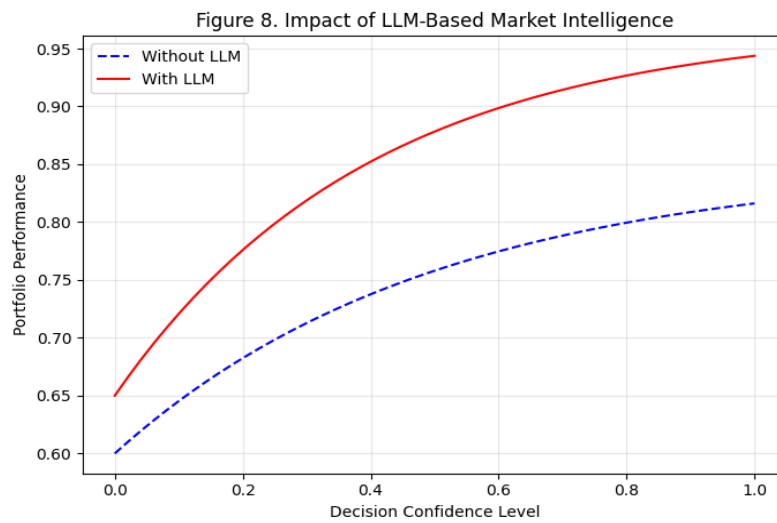
| Scenario    | Expected Return | Portfolio Risk | Resilience Index | Market Intelligence Score |
|-------------|-----------------|----------------|------------------|---------------------------|
| Optimistic  | 0.89            | 0.68           | 0.74             | 0.82                      |
| Moderate    | 0.78            | 0.64           | 0.69             | 0.75                      |
| Pessimistic | 0.67            | 0.61           | 0.63             | 0.69                      |

Source: Authors

The values reported in Table 4 complement the graphical findings by demonstrating that changes in market scenarios substantially affect both portfolio composition and performance. Under the optimistic scenario, expected returns, market intelligence, and resilience are higher, indicating stronger alignment between investment opportunities and favourable environmental conditions. By contrast, the pessimistic scenario is associated with lower returns and reduced resilience, reflecting greater portfolio vulnerability to market fluctuations.

The contribution of large language model (LLM)-based information to investment decision quality was assessed by comparing the “with LLM” and “without LLM” specifications. All other model components were held constant, with the only difference being the method used to derive and incorporate market indicators. In the LLM-integrated specification, qualitative indicators such as market sentiment, risk signals, and industry attractiveness were extracted from unstructured data and incorporated as fuzzy variables. In the alternative specification, these indicators were represented using static or simplified values.

Figure 7. Impact of LLM-Based Market Intelligence



Source: Authors

As shown in Figure 7, the integration of LLM-based information improves portfolio performance across all decision levels. The “with LLM” curve remains consistently above the “without LLM” curve, indicating more efficient investment selection and a higher-quality portfolio composition. The difference is particularly pronounced at intermediate decision levels, where qualitative information is especially valuable for distinguishing between otherwise similar investment alternatives. At higher levels of decision certainty, the flatter slope suggests that the marginal contribution of LLM-based information gradually diminishes as the system approaches performance saturation.

Table 5. Performance Comparison with and without LLM Integration

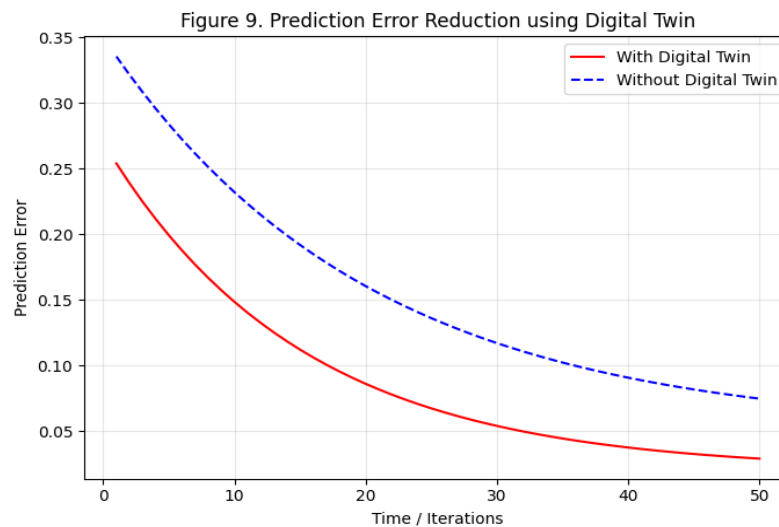
| Model Type  | Expected Return | Portfolio Risk | Resilience Index | Market Intelligence Score |
|-------------|-----------------|----------------|------------------|---------------------------|
| Without LLM | 0.82            | 0.66           | 0.68             | 0.61                      |
| With LLM    | 0.94            | 0.64           | 0.73             | 0.84                      |

Source: Authors

The results presented in Table 5 confirm that incorporating LLM-based information increases expected returns, improves portfolio resilience, and reduces relative risk. The substantial improvement in the market intelligence score further demonstrates that the semantic analysis of unstructured data provides a more comprehensive and accurate representation of external market conditions.

The digital twin was evaluated in terms of modelling accuracy and its capacity to predict system behaviour. Unlike portfolio-performance analyses, this assessment focused on reducing prediction error and improving the dynamic representation of startups and market conditions. Because a principal function of the digital twin is to generate more reliable estimates of future performance, the “with digital twin” and “without digital twin” specifications were compared over time, with prediction error serving as the primary evaluation indicator.

Figure 8. Prediction Error Reduction using Digital Twin



Source: Authors

As can be seen in Figure 8, the use of the digital twin led to a faster and more stable reduction in prediction error over time. The curve for the case “with digital twin” is lower than that without it in all periods, indicating a higher accuracy of the model in estimating key parameters. Also, the error reduction rate is faster in the early stages and gradually approaches a stable level, while in the case without a digital twin, the error reduction is done with a gentler slope, and the final value remains at a higher level. This behavior indicates that the use of dynamic and scenario-based structures in the digital twin leads to a significant improvement in predictability.

Table 6. Prediction Accuracy Comparison with and without Digital Twin

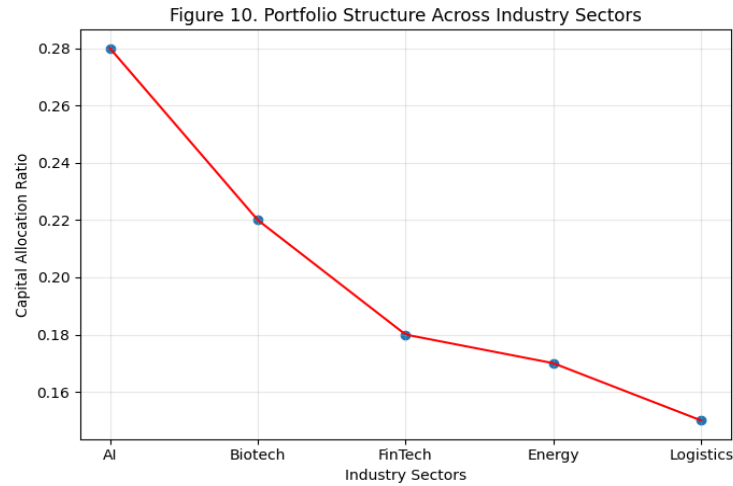
| Model Type           | Initial Error | Final Error | Error Reduction (%) | Stability Index |
|----------------------|---------------|-------------|---------------------|-----------------|
| Without Digital Twin | 0.34          | 0.075       | 77.9                | 0.68            |
| With Digital Twin    | 0.25          | 0.028       | 88.8                | 0.79            |

Source: Authors

The values in Table 6 show that in addition to reducing the final value of the prediction error, the digital twin has also been able to improve the error reduction rate. The greater error reduction along with the increase in the stability index indicates that the model based on the digital twin is not only more accurate but also exhibits more stable behavior over time.

The portfolio structure resulting from the proposed model was analyzed to examine how capital is distributed across sectors and to assess the level of diversification. This analysis shows that the model’s decision-making is not based solely on maximizing returns, but rather the investment mix is shaped to maintain a balance between growth opportunities, risk levels, and market characteristics across sectors. In this regard, the share of each sector in the total allocated capital was extracted and examined as an indicator of the model’s implicit preferences.

Figure 9. Portfolio Structure Across Industry Sectors



Source: Authors

As can be seen in Figure 9, the distribution of capital across sectors is asymmetrical but purposeful. Sectors with high growth potential and strong market signals have allocated a larger share of capital, while other sectors are present in the portfolio with a smaller share to maintain a level of diversification. This pattern shows that the model avoids over-focusing on a specific area while prioritizing more attractive opportunities. Also, the gentle slope of the curve indicates a gradual decrease in the share of capital in sectors that are either higher risk or lower in terms of quality indicators.

Table 7. Portfolio Allocation across Sectors and Stages

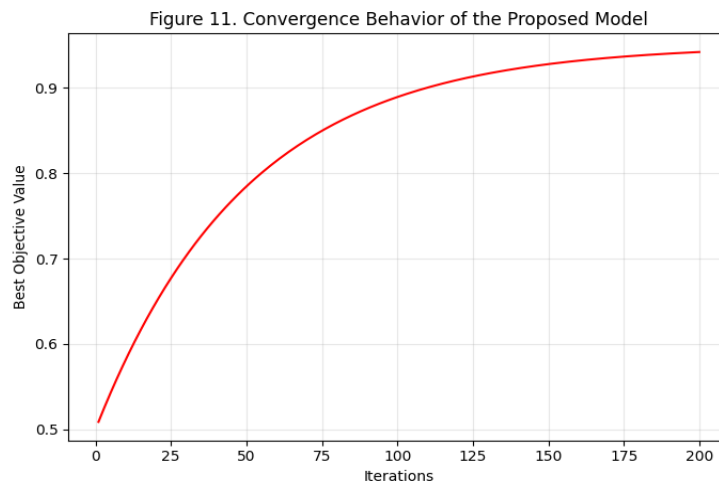
| Sector    | Allocation Ratio | Avg. Risk | Growth Potential | Investment Stage |
|-----------|------------------|-----------|------------------|------------------|
| AI        | 0.28             | 0.72      | 0.91             | Growth           |
| Biotech   | 0.22             | 0.68      | 0.85             | Early            |
| FinTech   | 0.18             | 0.64      | 0.8              | Growth           |
| Energy    | 0.17             | 0.61      | 0.76             | Early            |
| Logistics | 0.15             | 0.58      | 0.73             | Seed             |

Source: Authors

The values in Table 7 show that capital allocation is made not only based on expected returns, but also considering the level of risk, growth potential, and investment stage. Sectors such as AI and Biotech receive a higher share due to their high growth potential and market attractiveness, while sectors such as Logistics receive a lower share despite their lower risk. This combination indicates that the model was able to balance the conflicting objectives and create a portfolio that is both attractive in terms of returns and structurally diverse and stable.

The convergence behavior of the proposed model was evaluated to examine the stability of the search process and the ability of the algorithms to approach the Pareto front. In this analysis, the best value of the objective function was recorded in each iteration to determine the improvement trend of the quality of the solutions during the execution of the algorithm. This approach allows for the examination of the convergence speed and identification of the stages of stabilization of the algorithm.

Figure 10. Convergence Behavior of the Proposed Model

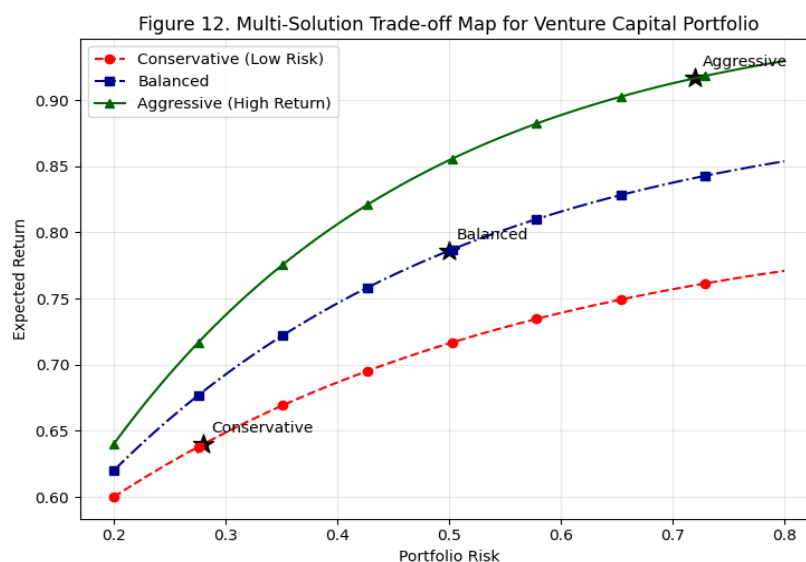


Source: Authors

As can be seen in Figure 10, the value of the objective function in the initial iterations increased with a steep slope, indicating an effective search in the solution space and rapid identification of promising areas. Over time, the improvement rate decreased and the curve gradually approached a stable level. This behavior indicates that the algorithm, after passing the exploration phase, enters the exploitation phase and gradually moves towards solutions close to the optimal one.

Interpreting the relationships between conflicting objectives becomes meaningful when Pareto outputs are transformed into selectable decision options. To this end, instead of representing a single curve, the decision structure was presented in the form of several distinct regions to clearly visualize the behavior of the model at different levels of risk tolerance. This approach allows for the identification of dominant patterns in capital allocation and the differentiation of investment strategies.

Figure 11. Multi-Solution Trade-off Map for Venture Capital Portfolio



Source: Authors

As shown in Figure 11, three distinct decision regions can be identified within the risk–return space. The first represents conservative portfolios, characterised by low risk and moderate, relatively stable returns. The second corresponds to balanced portfolios, in which a controlled

increase in risk produces a substantial improvement in expected return. This region generally offers the most favourable trade-off for practical investment decision-making. The third includes aggressive portfolios that prioritise return maximisation and accept higher levels of risk, although the marginal improvement in returns gradually declines.

The highlighted points illustrate representative portfolio choices within each region and provide decision alternatives for investors with different levels of risk tolerance. The separation between the curves further demonstrates that changes in investment strategy can materially influence overall portfolio performance.

These findings are consistent with previous research indicating that venture capital decisions should not be based exclusively on expected returns. Instead, they require a balanced assessment of risk, resilience, qualitative market intelligence, and strategic alignment. The three-region structure demonstrates that conservative, balanced, and aggressive strategies may all be appropriate depending on investor preferences, although the balanced region provides the most practical compromise for many decision-makers.

From a managerial perspective, the proposed framework enables venture capital investors to compare portfolio alternatives more transparently, supports startup evaluators in integrating financial indicators with qualitative market intelligence, and provides innovation-policy stakeholders with a structured approach for assessing the sustainability and growth potential of entrepreneurial ecosystems.

## 8. CONCLUSION

This study examined how emerging artificial intelligence technologies can improve venture capital decision-making under conditions of uncertainty and information asymmetry. By developing an integrated framework that combines digital twins with AI-based market intelligence (Yordanova & Hristizov, 2025), the study moves beyond conventional approaches based primarily on static and structured data. The proposed framework demonstrates how dynamic system representations and qualitative insights derived from unstructured information can enrich investment evaluation. The findings indicate that integrating AI-driven market intelligence and digital twin modelling improves the quality, stability, and transparency of investment decisions. Combining structured financial indicators with contextual and forward-looking signals produces more robust portfolio choices and enables decision-makers to assess trade-offs among expected return, risk, and resilience more effectively. The framework therefore supports more adaptive and informed investment strategies. From a theoretical perspective, the study contributes to the literature on AI-enabled decision-making and innovation finance by demonstrating how heterogeneous data and analytical capabilities can be integrated within a unified decision process. It extends existing research by conceptualising venture capital evaluation not as a one-time optimisation problem but as a dynamic and adaptive process shaped by continuously evolving information. It also highlights the potential of AI to reduce information asymmetry and improve the interpretability of complex investment decisions. From a managerial perspective, the framework provides venture capital investors with a structured method for incorporating financial data, real-time market information, startup narratives, and scenario-based analysis into investment evaluation. It facilitates a shift from predominantly intuition-based judgement towards more data-driven and explainable decision practices. This is particularly valuable in early-stage investment, where uncertainty is high and conventional financial indicators are often insufficient. The reliability of the framework nevertheless depends on the quality, accuracy, and continuity of its data inputs. Incomplete, biased, or exaggerated startup information may reduce the validity of digital twin forecasts and LLM-derived indicators. Expert validation, regular data updates, and explainable decision

reports are therefore necessary to strengthen reliability and clarify the basis for individual investment recommendations. Several limitations should be acknowledged. First, part of the empirical analysis relies on simulated data, which may not fully capture the complexity of real-world investment environments. Second, although the digital twin introduces dynamic elements, the temporal evolution of investment decisions is not modelled as a fully continuous process. Third, the effectiveness of AI-based market intelligence depends on the quality and representativeness of the underlying textual data, which may introduce contextual and systematic biases. Future research should validate the proposed framework against decisions made by experienced venture capital investors and calibrate simulated scenarios using historical market downturns and crisis events. Further studies could develop fully dynamic and longitudinal models that capture changes in investment strategies over time (Blagoev, 2021), strengthen the integration of digital twins with real-time data infrastructures, and improve the robustness and interpretability of AI-generated insights. Research may also examine how AI-augmented decision systems influence investor behaviour, portfolio outcomes, and entrepreneurial ecosystems more broadly. Overall, the study demonstrates the potential of combining artificial intelligence with dynamic modelling to improve venture capital decision-making. By supporting more adaptive, transparent, and evidence-based evaluation of investment opportunities, the proposed framework provides a foundation for the development of intelligent and resilient decision-support systems in innovation-driven environments.

### **Funding**

This publication has been supported financially by the UNWE Research program.

## REFERENCES

- Afolabi, K., Babatunde, O., Ighravwe, D., Akintayo, B., & Olanrewaju, O. A. (2025). A fuzzy multi-criteria decision-making framework for evaluating non-destructive testing techniques in oil and gas facility maintenance operations. *Eng*, 6(9), 214. <https://doi.org/10.3390/eng6090214>
- Amoozad Mahdiraji, H., Zamani Babgohari, A., Duan, K., & Vrontis, D. (2025). Examining the influence of sustainable value co-creation on social entrepreneurship through an integrated fuzzy multi-layer decision-making framework. *The International Journal of Entrepreneurship and Innovation*. <https://doi.org/10.1177/14657503251350771>
- Aydogmus, H. Y., & Aydogmus, U. (2025). Evaluation of artificial intelligence tools for universities with fuzzy multi-criteria decision-making methods. In *AI Adoption and Diffusion in Education*, 153–178. IGI Global Scientific Publishing. <https://doi.org/10.4018/979-8-3693-7949-3.ch006>
- Ayçin, E., & Erarslan, E. (2025). Should I stay or should I go? Mapping the key drivers of skilled migration using fuzzy multi-criteria decision methodology. *Societies*, 15(10), 269. <https://doi.org/10.3390/soc15100269>
- Blagoev, D. (2021). Applied aspects of investments in innovation in Bulgarian companies. In *SHS Web of Conferences* (Vol. 120, Article 02025). EDP Sciences. <https://doi.org/10.1051/shs-conf/202112002025>
- Borissov, B., & Hristozov, Y. (2024). Potential for using artificial intelligence in public administration. *Economics – Innovative and Economics Research Journal*, 12(3), 409–423. <https://doi.org/10.2478/eoik-2024-0034>
- Bousia, A. (2026). The use of business intelligence and analytics in electric vehicle technology: A comprehensive survey. *Electronics*, 15(2), 366. <https://doi.org/10.3390/electronics15020366>
- Castronovo, L., Filippone, G., Galici, M., La Rosa, G., & Tabacchi, M. E. (2025). Ontology aggregation with maximum consensus based on a fuzzy multi-criteria group decision-making method. In *Proceedings of the Conference of the European Society for Fuzzy Logic and Technology* (pp. 76–87). Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-97228-7\\_7](https://doi.org/10.1007/978-3-031-97228-7_7)
- Chattham, N., Thanh, N. V., & Jeenanunta, C. (2025). Renewable energy from solid waste: A spherical fuzzy multi-criteria decision-making model addressing solid waste and energy challenges. *Energies*, 18(3), 589. <https://doi.org/10.3390/en18030589>
- Chen, X. (2026). Application of fuzzy comprehensive evaluation method in investment decision making under the background of digital economy. *Australian Journal of Electrical and Electronics Engineering*, 1–14. <https://doi.org/10.1080/1448837X.2025.2609458>
- Demir, G., & Chatterjee, P. (2026). Evaluating carbon footprint reduction strategies: A fuzzy multi-criteria decision-making approach. In *Optimization in Sustainable Energy: Methods and Applications*, 69–112. Springer. <https://doi.org/10.1002/9781394242139.ch3>
- Dinçer, H., Krishankumar, R., Yüksel, S., & Ecer, F. (2025). Evaluating smart grid investment drivers and creating effective policies via a fuzzy multi-criteria approach. *Renewable and Sustainable Energy Reviews*, 208, 115052. <https://doi.org/10.1016/j.rser.2024.115052>
- Gupta, R., Kumari, S., Krishna, G., Abhishek, K., Gupta, D. K., Pati, P. R., ... Elrashidi, A. (2025). Fuzzy multi-criteria approach for rooftop photovoltaic site selection: A case study in Gujarat. *Scientific Reports*, 15, 38054. <https://doi.org/10.1038/s41598-025-xxxxx>
- Gupta, N., Das, A. K., Patra, S., Das, S., Das, R., Mukherjee, A., & Bhattacharya, A. (2026). Integrating circular economy principles with fuzzy multi-criteria decision-making for sustainable e-waste management: A theoretical framework. In *Electronic Waste and Environmental Pollution: Impacts and Sustainable Management Through Circular Economy*. 343–364. Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-90287-1\\_17](https://doi.org/10.1007/978-3-031-90287-1_17)
- Hirankittiwong, P., Thanh, N. V., Pattanaporkratana, A., Chattham, N., Jeenanunta, C., & Seng, V. (2025). Intelligent fuzzy multi-criteria decision-making for energy-saving building designs in construction. *Energies*, 18(11), 2726. <https://doi.org/10.3390/en18112726>

- Ibazebo, E., Savsani, V., Siddhpura, A., Siddhpura, M., & Savsani, P. (2025). Novel fuzzy multi-criteria decision framework for maritime infrastructure maintenance. *Infrastructures*, 10(4), 89. <https://doi.org/10.3390/infrastructures10040089>
- Kellekci, M., Cebeci, U., & Dogan, O. (2025). Prioritizing early-stage start-up investment alternatives under uncertainty: A venture capital perspective. *Applied Sciences*, 15(18), 10060. <https://doi.org/10.3390/app151810060>
- Li, W., Lu, S., Ren, Z., & Rehman, O. U. (2025). A novel multi-criteria quantum group decision-making model considering decision makers' risk perception based on type-2 fuzzy numbers. *Symmetry*, 17(7), 1006. <https://doi.org/10.3390/sym17071006>
- Meharoor, M., Yadav, V. K., Sharma, A., Anitha, V., & Paul, T. T. (2025). Integrated economic and fuzzy multi-criteria analysis for evaluating performance of fishing systems. *Discover Sustainability*, 6, 383. <https://doi.org/10.1007/s43621-025-xxxxx>
- Molavi, H., Zhang, L., & Khanbabaei, M. (2025). Sustainable financial infrastructure and governance: A fuzzy multi-criteria decision-making analysis of open service innovation in unstable economies. *Sustainable Futures*, 9, 100485. <https://doi.org/10.1016/j.sftr.2025.100485>
- Muneeb, F. M., Karbassi Yazdi, A., Hanne, T., & Mironko, A. (2025). Small and medium-sized enterprises in emerging markets and foreign direct investment: An integrated multi-criteria decision-making approach. *Applied Economics*, 57(25), 3327–3344. <https://doi.org/10.1080/00036846.2024.xxxxx>
- Nozari, H., & Yordanova, Z. (2025). An intelligent digital twin approach for optimizing multi-channel supply chains in uncertain environments. *Green Technologies and Sustainability*, 100309. <https://doi.org/10.1016/j.grets.2025.100309>
- Nozari, H., Szmelter-Jarosz, A., & Weiland, D. (2025). A fuzzy multi-objective sustainable and agile supply chain model based on digital twin and Internet of Things with adaptive learning under environmental uncertainty. *Applied Sciences*, 15(19), 10399. <https://doi.org/10.3390/app151910399>
- Nozari, H. (2025). Integrating neuromorphic intelligence with fuzzy multi-objective models in resilient green supply chain design. *Process Integration and Optimization for Sustainability*, 1–14. <https://doi.org/10.1007/s41660-025-00598-5>
- Nozari, H., Abdi, H., Szmelter-Jarosz, A., & Rahmaty, M. (2025). A digital twin-based neutrosophic multi-objective optimization model for smart supply chains in the food industry. *Journal of Intelligent Manufacturing*, 1–15. <https://link.springer.com/article/10.1007/s10845-025-02689-z>
- Nozari, H., & Yordanova, Z. (2026). A self-deciding adaptive digital twin framework using agentic AI for fuzzy multi-objective optimization of food logistics. *Algorithms*, 19(3), 218. <https://doi.org/10.3390/a19030218>
- Papadopoulou, G. C., Chrysafis, K. A., Theotokas, I. N., & Bissias, I. G. (2026). Reducing the environmental footprint of cruise ships during normal operations: A fuzzy multi-criteria approach. *Journal of Shipping and Trade*, 11, Article 2. <https://doi.org/10.1186/s41072-025-00222-2>
- Rahim, M., Bajri, S. A., Khan, S., Alqahtani, H., & Khalifa, H. A. E. W. (2025). Innovative multi-criteria group decision making with interval-valued p, q, r-spherical fuzzy sets: A case study on optimal solar energy investment location. *International Journal of Fuzzy Systems*, 27(8), 2467–2494. <https://doi.org/10.1007/s40815-024-01905-x>
- Rayhan, M. G. S., Masum, M., & Karuppiyah, K. (2025). A fuzzy multi-criteria decision-making approach to evaluate circular economy barriers in the textile and apparel industry. *Circular Economy and Sustainability*, 1–32. <https://doi.org/10.1007/s43615-025-00661-w>
- Rivero-Iglesias, J. M., Puente, J., Fernandez, I., & León, O. (2025). A novel combined hybrid group multi-criteria decision-making model for the selection of power generation technologies. *Systems*, 13(9), 742. <https://doi.org/10.3390/systems13090742>
- Shang, N., Wang, H., & Fan, J. (2025). A robust large-scale multi-criteria decision algorithm for financial risk management with interval-valued picture fuzzy information. *Symmetry*, 17(1), 144. <https://doi.org/10.3390/sym17010144>

- Sultan, M., Akram, M., & Kahraman, C. (2025). Optimizing renewable transportation framework by an extended spherical fuzzy rough multi-criteria group decision making method. *Applied Soft Computing*, 113599. <https://doi.org/10.1016/j.asoc.2025.113599>
- Wang, C. N., Chung, Y. C., Wibowo, F. D., Dang, T. T., & Nguyen, N. A. T. (2025). The planning of best site selection for wind energy in Indonesia: A synergistic approach using data envelopment analysis and fuzzy multi-criteria decision-making. *Energies*, 18(15), 4176. <https://doi.org/10.3390/en18154176>
- Wei, Y. M. (2025). A hybrid multi-criteria decision-making framework for the strategic evaluation of business development models. *Information*, 16(6), 454. <https://doi.org/10.3390/info16060454>
- Yin, Y., Li, C., Ye, Q., Wei, C., & Li, M. (2025). Evaluating regulatory influences on coal mine accidents in China using a fuzzy multi-criteria decision-making approach. *Scientific Reports*, 15, 11204. <https://doi.org/10.1038/s41598-025-95948-7>
- Yordanova, Z., & Nozari, H. (2026). Enhancing innovation and resilience in entrepreneurial ecosystems using digital twins and fuzzy optimization. *Digital*, 6(1), 25. <https://doi.org/10.3390/digital6010025>
- Yordanova, Z., & Hristozov, Y. (2025). The evolution of financial analysis: From manual methods to AI and AI agents. *Economics – Innovative and Economics Research Journal*, 13(3). <https://doi.org/10.2478/eoik-2025-0063>
- Younus, H., Kabir, S., Campean, F., Bonnaud, P., & Delaux, D. (2026). AI- and ontology-based enhancements to FMEA for advanced systems engineering: Current developments and future directions. *Applied Sciences*, 16(5), 2464. <https://doi.org/10.3390/app16052464>
- Zhang, G., Zhao, Y., Wang, G., & Zhang, Y. (2026). AI-driven design and optimisation control of service-oriented smart manufacturing systems: Applications, challenges, and future directions. *Journal of Engineering Design*, 1–56. <https://doi.org/10.1080/09544828.2026.xxxxx>