

AN INTEGRATED AHP–TOPSIS–KNAPSACK–LLM HYBRID MODEL FOR
PROJECT PORTFOLIO OPTIMIZATION

Shamsiyev Azizbek Rafiq ugli

Master's student, Digital Economy and Agrotechnologies University,
Tashkent, UzbekistanE-mail: azizbekshamsiyev946@gmail.com | Tel.: +998 90 053 04 08<https://doi.org/10.5281/zenodo.20144717>

Abstract. This paper proposes an integrated four-component hybrid pipeline for project portfolio optimization that combines the Analytic Hierarchy Process (AHP) for criterion weighting, TOPSIS for alternative ranking, 0/1 Knapsack dynamic programming for budget-constrained optimal selection, and GPT-4o based large-language-model (LLM) scoring for quantifying qualitative project documents. The mathematical apparatus, overall computational complexity $O(n \cdot m + n \cdot B)$, and Design Science Research based experimental protocol are formalized. Validation is planned on a synthetic dataset of 20 projects across 30 replications against three baselines (pure NPV greedy, greedy heuristic, and pure MCDM). Four falsifiable hypotheses (H1–H4) are tested using paired t-tests, the Wilcoxon signed-rank test, and Monte Carlo sensitivity analysis. The novelty of the work lies in being the first fully integrated AHP+TOPSIS+0/1 Knapsack+LLM pipeline reported in the literature, with direct applicability to the IT initiatives planned under the "Digital Uzbekistan 2030" strategy.

Keywords: project portfolio, optimization, AHP, TOPSIS, 0/1 Knapsack, large language model, GPT-4o, multi-criteria decision making, design science research.

Annotatsiya. Mazkur maqolada loyiha portfelin optimallashtirish uchun to'rt komponentli integratsiyalashgan gibrid pipeline taklif etiladi: mezon og'irliklarini topish uchun AHP, alternativlarni tartiblash uchun TOPSIS, byudjet konstrainti ostida optimal tanlash uchun 0/1 Knapsack dynamic programming va sifatli loyiha hujjatlaridan miqdoriy ball olish uchun GPT-4o asosidagi LLM-scoring. Pipeline-ning matematik apparati, umumiy hisoblash murakkabligi $O(n \cdot m + n \cdot B)$ va Design Science Research asosidagi eksperimental protokol rasmiylashtirilgan. Validatsiya 20 ta sintetik loyihada 30 marta uchta baseline metodga qarshi rejalashtirilgan. To'rtta falsifikatsiyalanuvchi gipoteza (H1–H4) paired t-test, Wilcoxon signed-rank test va Monte Carlo sensitivity tahlili yordamida tekshiriladi. Ishning yangiligi adabiyotda birinchi marta to'liq integratsiyalashgan AHP+TOPSIS+0/1 Knapsack+LLM pipeline-ini taqdim etishidir va u "Raqamli O'zbekiston-2030" strategiyasi doirasidagi IT tashabbuslarga to'g'ridan-to'g'ri qo'llanilishi mumkin.

Kalit so'zlar: loyiha portfeli, optimallashtirish, AHP, TOPSIS, 0/1 Knapsack, katta til modeli, GPT-4o, ko'p mezonli qaror qabul qilish, loyihaviy ilmiy tadqiqot.

Аннотация. В статье предложен интегрированный четырёхкомпонентный гибридный конвейер для оптимизации портфеля проектов, объединяющий метод анализа иерархий (AHP) для расчёта весов критериев, TOPSIS для ранжирования альтернатив, динамическое программирование 0/1 Knapsack для оптимального выбора при бюджетных ограничениях и LLM-оценивание на базе GPT-4o для количественной обработки качественных проектных документов. Формализованы математический аппарат, общая вычислительная сложность $O(n \cdot m + n \cdot B)$ и экспериментальный протокол на основе Design Science Research. Валидация

запланирована на синтетическом наборе из 20 проектов с 30 повторениями против трёх базовых методов. Четыре опровергаемые гипотезы (H1–H4) проверяются парным t-тестом, критерием Уилкоксона и анализом чувствительности Монте-Карло. Новизна работы заключается в представлении первого полностью интегрированного AHP+TOPSIS+0/1 Knapsack+LLM конвейера и его применимости к ИТ-инициативам стратегии "Цифровой Узбекистан-2030".

Ключевые слова: портфель проектов, оптимизация, AHP, TOPSIS, 0/1 Knapsack, большая языковая модель, GPT-4o, многокритериальное принятие решений, проектная наука.

1. Introduction

Modern organizations increasingly run dozens of IT projects in parallel under tight budget and resource constraints. The Standish Group CHAOS Report consistently shows that only about one third of IT projects are completed on time, on budget, and with the expected functionality, while the rest fail partially or completely (Standish Group, 2020). Mathematically, selecting an optimal subset from n candidate projects is NP-hard, with 2^n possible combinations. The decision is further complicated by well-documented cognitive biases — loss aversion, anchoring, and the availability heuristic — formalized in Prospect Theory (Kahneman & Tversky, 1979).

Recent advances in large language models (LLMs) based on the Transformer architecture (Vaswani et al., 2017), few-shot learning (Brown et al., 2020), and chain-of-thought prompting (Wei et al., 2022) open new opportunities for processing qualitative project documents. Empirical evaluations show that LLMs can already score above 82% on Project Management Institute certification questions (Karnouskos, 2024) and provide reproducible decision support for research project evaluation (Ma et al., 2024). However, no existing study has integrated AHP, TOPSIS, 0/1 Knapsack dynamic programming, and LLM scoring into a single mathematically formalized pipeline with rigorous statistical validation. This paper closes that gap and outlines applicability to the IT portfolio of the "Digital Uzbekistan 2030" strategy.

2. Theoretical foundations

Five theoretical streams underpin the proposed model. (i) Modern portfolio theory (Markowitz, 1952) formalizes the mean–variance trade-off but does not transfer directly to project portfolios due to illiquidity and non-stationary correlations. (ii) Project portfolio management frameworks (Archer & Ghasemzadeh, 1999; Cooper, Edgett, & Kleinschmidt, 2001; Project Management Institute, 2017) provide the staged process within which the model operates. (iii) Multi-criteria decision making (MCDM) supplies the ranking apparatus: AHP via pairwise comparison matrices and eigenvectors (Saaty, 1980), TOPSIS via Euclidean distances to ideal solutions (Hwang & Yoon, 1981; Opricovic & Tzeng, 2004), and BWM as a parsimonious alternative (Rezaei, 2015). (iv) Optimization theory contributes 0/1 Knapsack dynamic programming (Bellman, 1957; Martello & Toth, 1990) and multi-objective evolutionary algorithms (Deb et al., 2002). (v) Artificial-intelligence theory supplies the LLM component (Devlin et al., 2019; Vaswani et al., 2017; Brown et al., 2020; Wei et al., 2022). The methodology follows the Design Science Research paradigm (Hevner et al., 2004), with hypotheses formulated under Popper's (1959) falsifiability principle.

3. The hybrid model

AHP. A pairwise comparison matrix $A=[a_{ij}]$ on the 1–9 Saaty scale yields the weight vector w from the principal eigenvector $A \cdot w = \lambda_{\max} \cdot w$. Consistency is verified through $CI=(\lambda_{\max}-n)/(n-1)$ and $CR=CI/RI$, requiring $CR<0.10$ (Saaty, 1980). For the five criteria used

here (financial efficiency, risk, strategic fit, technical complexity, team capability) expert elicitation produced $w=(0.25, 0.20, 0.20, 0.20, 0.15)$ with $CR=0.0045$.

TOPSIS. After vector normalization $r_{ij} = x_{ij} / \sqrt{(\sum x_{ij}^2)}$ and weighting $v_{ij} = w_j \cdot r_{ij}$, the closeness coefficient

$$C_i = D_i^- / (D_i^+ + D_i^-), \quad C_i \in [0, 1]$$

is computed for each alternative, with D_i^+ and D_i^- denoting Euclidean distances to the ideal A^* and anti-ideal A^- solutions (Hwang & Yoon, 1981).

0/1 Knapsack. Composite values $v_i = 100 \cdot C_i$ are passed into the budget-constrained selection problem

$$\max \sum_i v_i \cdot x_i, \quad \sum_i c_i \cdot x_i \leq B, \quad x_i \in \{0, 1\}$$

solved exactly by bottom-up dynamic programming in $O(n \cdot B)$, guaranteeing global optimality (Bellman, 1957; Martello & Toth, 1990).

LLM scoring. GPT-4o is prompted with a four-layer architecture (system role, PMI context, scoring task with chain-of-thought reasoning, and three few-shot examples). Stochasticity is reduced with temperature=0.2 and JSON-schema validation. The LLM outputs 0–100 scores for qualitative criteria that feed back into the TOPSIS decision matrix. Total complexity of the pipeline is $O(n \cdot m + n \cdot B)$, well within practical limits for portfolios of 50–500 projects.

4. Experimental design and expected results

The model is validated on a synthetic dataset of 20 projects organized into three risk–return families (A: high–high, 7 projects; B: medium–medium, 7 projects; C: low–low, 6 projects), with a global budget of $B=1,000,000$ USD. The experiment runs 30 replications and compares the proposed model (M3) against two baselines: pure NPV greedy (M1) and budget-aware greedy heuristic (M2). Four hypotheses are tested. H1 (effectiveness): M3 increases composite portfolio value by $\geq 10\%$ over M1 (paired t-test, $p < 0.05$). H2 (LLM consistency): GPT-4o re-scoring of the same project across 30 trials yields a coefficient of variation $CV < 10\%$. H3 (Knapsack efficiency): exact DP reduces portfolio risk by $\geq 10\%$ over the greedy baseline (Wilcoxon signed-rank, $p < 0.05$). H4 (stability): under $\pm 10\%$ perturbation of criterion weights, portfolio composition retains $\geq 70\%$ overlap (Monte Carlo, 100 iterations, Jaccard index). G*Power 3.1 confirms the required sample size $n \approx 26$ at $\alpha=0.05$, power=0.80, $d=0.5$; the chosen $n=30$ exceeds this threshold. Reproducibility is ensured by fixed random seed=42 and open-source release of all code.

5. Conclusion

The paper formalizes the first fully integrated AHP+TOPSIS+0/1 Knapsack+LLM pipeline for project portfolio optimization, with an explicit mathematical apparatus, polynomial complexity, and a rigorous Design Science Research validation plan. The contribution is both theoretical (closing the integration gap between MCDM, exact optimization, and modern LLMs) and practical (an open, reproducible tool directly aligned with Uzbekistan's national digital transformation agenda). Future work will extend validation to real corporate datasets, integrate Group AHP for multi-expert elicitation, incorporate explainable-AI components, and explore Uzbek-language LLM fine-tuning.

References

1. Archer, N. P., & Ghasemzadeh, F. (1999). An integrated framework for project portfolio selection. *International Journal of Project Management*, 17(4), 207–216. [https://doi.org/10.1016/S0263-7863\(98\)00032-5](https://doi.org/10.1016/S0263-7863(98)00032-5)

2. Bellman, R. (1957). *Dynamic programming*. Princeton, NJ: Princeton University Press.
3. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
4. Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (2001). *Portfolio management for new products* (2nd ed.). Cambridge, MA: Basic Books.
5. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197. <https://doi.org/10.1109/4235.996017>
6. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186.
7. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
8. Hwang, C. L., & Yoon, K. (1981). *Multiple attribute decision making: Methods and applications*. Berlin: Springer-Verlag. <https://doi.org/10.1007/978-3-642-48318-9>
9. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
10. Karnouskos, S. (2024). The relevance of large language models for project management. *IEEE Open Journal of the Industrial Electronics Society*, 5, 758–768. <https://doi.org/10.1109/OJIES.2024.3412222>
11. Ma, J., Zheng, Z., Zhu, P., & Liu, Z. (2024). A machine learning and large language model-integrated approach to research project evaluation. *Journal of Database Management*, 35(1), 1–22. <https://doi.org/10.4018/JDM.345400>
12. Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
13. Martello, S., & Toth, P. (1990). *Knapsack problems: Algorithms and computer implementations*. Chichester: John Wiley & Sons.
14. Opricovic, S., & Tzeng, G.-H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, 156(2), 445–455. [https://doi.org/10.1016/S0377-2217\(03\)00020-1](https://doi.org/10.1016/S0377-2217(03)00020-1)
15. Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.
16. Project Management Institute. (2017). *The standard for portfolio management* (4th ed.). Newtown Square, PA: PMI.
17. Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, 53, 49–57. <https://doi.org/10.1016/j.omega.2014.11.009>
18. Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. New York: McGraw-Hill.
19. Standish Group. (2020). *CHAOS Report 2020: Beyond infinity*. Boston, MA: The Standish Group International.
20. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
21. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35, 24824–24837.