Explainable Investment Recommendations with Knowledge Graph Path Reasoning: Insights from ALPHA10X

Amir Wolfensohn[1], Veronika Bogina[1], Issam Ibnouhsein[2], Didier Vila[2], Robert Marcus[2], Irad Ben-Gal[1,2], Noam Koenigstein[1,2]
[1]Tel Aviv University [2]ALPHA10X
Israel France

Abstract

We present an explainable investment recommender system developed on ALPHA10X's knowledge graph (KG) that leverages Knowledge-aware Path Attention (KPAR). The KG integrates diverse financial datasets, including Crunchbase, USPTO, executive profiles, industry classifications, and news signals, capturing complex investor–startup relationships. Our method employs KPAR to identify interpretable multi-hop reasoning paths between investors and startups. To extend interpretability beyond local explanations, we introduce a statistical framework for global interpretability, validating key investment signals—such as industry familiarity, angel investor influence, founder reputation, optimal co-investment levels, and geographic biases—through hypothesis testing. Experimental results demonstrate that our approach provides transparent and actionable insights while maintaining strong predictive accuracy, supporting high-stakes decision-making in venture capital markets.

CCS Concepts

• **Computing methodologies → Causal reasoning and diagnostics**; **Machine learning**; • **Information systems → Recommender systems**.

Keywords

Recommender Systems, Reasoning, Explainable Knowledge Graphs

ACM Reference Format:
Amir Wolfensohn[1], Veronika Bogina[1], Issam Ibnouhsein[2], Didier Vila[2],, Robert Marcus[2], Irad Ben-Gal[1,2], Noam Koenigstein[1,2] . 2018. Explainable Investment Recommendations with Knowledge Graph Path Reasoning: Insights from ALPHA10X. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 4 pages. https://doi.org/XXXXXXX.XXXXXXX

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

ALPHA10X is an AI-driven investment intelligence company that provides predictive analytics for venture capitalists, corporate strategists, and M&A advisors. Its *Nostradamus* platform leverages knowledge graph reasoning and deep learning to generate explainable, data-driven investment insights. A key challenge in this domain is delivering transparent, interpretable recommendations that align with investor decision-making. This research enhances ALPHA10X's system by uncovering global investment patterns through knowledge graph path analysis.

Investment recommendations involve high stakes and substantial financial risks, requiring clear explanations for effective decision-making. To address this, we propose an interpretable recommender framework that leverages ALPHA10X's investment knowledge graph (KG), which integrates data from Crunchbase[1], the U.S. Patent and Trademark Office (USPTO)[2], executive profiles, industry classifications, and news articles. Our model, the *Knowledge-aware Path Attentive Recommender (KPAR)* [4], aggregates interpretable paths connecting investors and startups through an attention mechanism.

To extend interpretability beyond local explanations, we introduce a statistical framework for *global interpretability*, analyzing KPAR's attention weights to identify key investment signals, including industry familiarity, angel investor influence, founder reputation, optimal co-investment strategies, and geographic preferences.

We demonstrate that our method uncovers robust global insights aligned with real-world investor behavior while maintaining strong predictive accuracy. Our approach offers a scientifically grounded, explainable framework for high-stakes investment decisions.

2 Related Work

Explainability in recommender systems increasingly leverages knowledge graphs (KGs), which explicitly represent entities and relationships, enabling structured reasoning over multi-hop connections [6, 8, 14, 23]. Path-based KG methods [4, 20, 22] have emerged as effective for generating interpretable recommendations by highlighting explicit reasoning paths [15, 16, 19, 21]. Recent advances incorporate attention mechanisms to identify influential paths, improving interpretability and model transparency [2, 4, 12].

Despite these advances, existing KG-based recommenders primarily focus on local explanations, offering insight only at the instance level without statistically grounded global patterns. Furthermore, few studies address high-stakes domains like venture capital investments, where interpretability is critical for decision-making. Notably, prior works on investment recommendation often

[1]https://www.crunchbase.com/
[2]https://www.uspto.gov/

lack fine-grained path-based reasoning [10] or rely on parameterized explanations that are not grounded in graph-based multi-hop analysis [9, 24].

3 Knowledge Graph Construction

Our approach leverages a subset of ALPHA10X's investment knowledge graph (KG), which integrates heterogeneous data sources capturing investor–startup relationships, industry context, and innovation signals. The KG includes:

- **Crunchbase data**: Organization attributes and funding history.
- **Executive and founder roles**: 1.4M person–organization links across 853K entities.
- **Industry categories**: 2.16M company–industry pairs spanning 152 sectors.
- **Scientific and patent concepts**: 126K organizations linked to 40K technologies.
- **News-derived relationships**: Parsed from press releases to capture real-time signals.

The final KG comprises 43K entities and 1M triplets, structured across 7 entity types (e.g., Investor, Company, Person, Industry) and 5 relation types (`invested_in`, `executive_in`, `has_concept`, `located_in`, `industry`). Each edge encodes a relational fact relevant to investment reasoning.

4 Path-Based Investment Reasoning with KPAR

We employ *Knowledge-aware Path Attentive Recommender (KPAR)* [4] to generate investment recommendations by leveraging a knowledge graph (KG) that models relationships between investors and startups. Given an investor u and a target company i, KPAR samples multiple relational paths $P(u, i) = \{p_1, p_2, \ldots, p_k\}$ connecting these entities within the KG.

Path Representation and Encoding. Each path $p \in P(u, i)$ is encoded into a vector representation using a learned path encoder:

$$\mathbf{v}_p = f_\theta(p),$$

where f_θ is the path encoding function parameterized by θ. For example, consider the path:

`Investor Alice → (invested_in) Startup Z → (founder) Bob → (founder_of) BioTechX`

This path captures a meaningful connection: Alice's investment in Startup Z, co-founded by Bob (who also founded BioTechX), signals familiarity and trust. KPAR represents such multi-hop reasoning in a latent space, effectively encoding both semantic and structural information.

Cross-Attention Mechanism. KPAR aggregates these path representations using a cross-attention mechanism that prioritizes paths based on their relevance to the investor-company pair. The aggregated representation is computed as:

$$\mathbf{x}_{ui} = \sum_{p \in P(u,i)} \alpha_p \mathbf{v}_p,$$

where α_p denotes the attention weight for path p, reflecting its contribution to the final recommendation. These weights are calculated through a softmax operation over the dot products of the path vectors and a query vector $\mathbf{q}_{ui}$:

$$\alpha_p = \frac{\exp(\mathbf{q}_{ui}^\top \mathbf{W}_k \mathbf{v}_p)}{\sum_{p' \in P(u,i)} \exp(\mathbf{q}_{ui}^\top \mathbf{W}_k \mathbf{v}_{p'})},$$

where $\mathbf{W}_k$ is a learnable projection matrix and $\mathbf{q}_{ui}$ is derived by concatenating the investor and company embeddings.

Prediction Head. The final affinity score, representing the likelihood of investment, is computed by passing the aggregated path representation through a prediction layer:

$$\hat{y}_{ui} = \sigma(\mathbf{w}^\top \mathbf{x}_{ui}),$$

where $\sigma(\cdot)$ is the sigmoid activation function and $\mathbf{w}$ is a learnable parameter vector. This score estimates the compatibility between the investor and the startup, leveraging structured reasoning paths in the KG.

Interpretability through Attention Weights. KPAR provides interpretability through its attention mechanism. Paths with higher weights are deemed more influential for the recommendation. In the example path above, if the attention weight α_p is high, it suggests that Alice's prior connection with Bob—who also founded BioTechX—is a major driver of the recommendation. Our analysis extends KPAR's interpretability from local explanations to *global interpretability* by statistically analyzing these weights to extract business insights, such as industry familiarity, angel investor influence, and geographic preferences.

Global Interpretability via Statistical Validation. While traditional attention mechanisms offer local interpretability, our approach introduces a rigorous framework for *global interpretability.* We statistically validate the influence of attention-weighted paths across large-scale investor-startup pairs, turning attention scores into actionable business insights.

We aggregate attention weights by semantic groupings—such as *angel-backed* vs. *VC-backed* paths, *founder-driven* vs. *industry-driven* paths, and varying levels of prior VC involvement. Statistical hypothesis testing is then performed to identify significant investment signals. For each hypothesis A versus a baseline B, we test:

$$H_0 : \mu_{\alpha,A} = \mu_{\alpha,B}, \quad H_1 : \mu_{\alpha,A} \neq \mu_{\alpha,B}$$

where $\mu_{\alpha,A}$ and $\mu_{\alpha,B}$ represent the mean attention weights for paths in patterns A and B, respectively. We apply robust statistical tests (e.g., t-tests, Mann-Whitney U tests) to validate these differences, ensuring that observed behaviors reflect genuine, data-driven investment preferences.

Our framework elevates KPAR's interpretability from isolated path explanations to statistically validated *global investment signals*, revealing critical patterns in investor behavior. In the following Results section, we demonstrate how these signals—such as angel investor influence, founder reputation, and optimal co-investment saturation—align with real-world investment strategies.

5 Results and Key Insights

We evaluate KPAR's effectiveness both in terms of **predictive accuracy** and its ability to extract **global interpretability signals** from path attention weights. Our experimental setup involved an 80/20 random split of the dataset, with 86,541 investor–company interactions for training and the remaining 20% held out for testing. All hyperparameters, including embedding dimensions, hidden sizes, dropout rates, and learning rates, were fixed across experiments.

5.1 Predictive Accuracy and Baseline Comparison

While the primary focus of our work is explainability, strong recommendation accuracy remains crucial for practical adoption. We compared KPAR against three competitive baselines:

- **Matrix Factorization (MF)** - Collaborative filtering-based.
- **Popularity-based Ranking** - Prioritizes historically popular startups.
- **PathLength Heuristic** - Favors startups with shorter paths to investors in the KG.

We evaluated performance using standard top-K metrics: Hit Rate, nDCG, Precision, and Recall. Figure 1 illustrates that KPAR consistently outperforms the baselines, particularly at lower ranks critical for decision-making (e.g., nDCG@5 of 0.94 vs. 0.86 for MF). These results come to demonstrate that emphasizing interpretability through path reasoning did not compromise predictive accuracy.

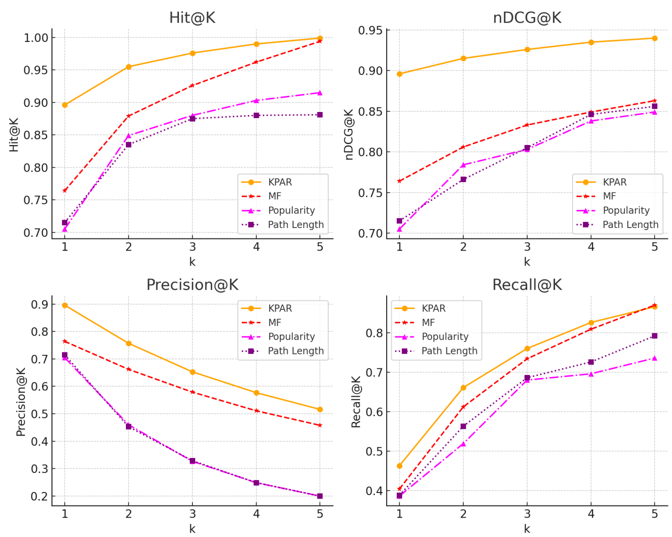

Figure 1: Top-K recommendation accuracy comparison. KPAR achieves superior performance relative to MF, Popularity, and PathLength, highlighting its capability to maintain interpretability alongside accuracy.

5.2 Global Investment Signals from Path Attention

By statistically analyzing KPAR's attention weights, we identify key global investment patterns, each rigorously validated through hypothesis testing ($p < 0.05$).

Industry Familiarity as an Investment Signal. Investors show a preference for startups within industries where they have prior experience, as evidenced by significantly higher path attention scores ($p < 0.05$). However, industry familiarity alone, without supporting signals, does not fully predict investment choices ($p = 0.717$). This finding suggests that industry knowledge is valuable but typically combined with other factors.

Examples: SoftBank's substantial investments in ride-hailing companies like Uber, Didi, and Grab demonstrate a strategic focus on the transportation sector [13]. Similarly, Andreessen Horowitz's early investment in Coinbase and subsequent emphasis on crypto startups highlight the role of industry familiarity in guiding investment decisions [1].

Angel Investors as Early-Stage Quality Signals. Paths involving angel investors consistently show higher average attention weights compared to those involving only VCs (mean: 0.67 vs. 0.53, $p < 0.001$). This confirms that angels act as early-stage validators, signaling quality and reducing uncertainty for later investors.

Examples: Peter Thiel's $500k seed investment in Facebook served as a strong quality signal, leading to Accel's $12.7M Series A investment [3]. Google similarly benefited from early backing by Andy Bechtolsheim and Jeff Bezos, bolstering its credibility [11].

Founder Reputation as a Key Investment Signal. Founder involvement emerged as one of the strongest predictors of investment, with paths involving founders achieving significantly elevated attention scores ($p < 10^{-5}$). This highlights investor reliance on founder credibility and track records.

Examples: SoftBank's $4.4B commitment to WeWork was heavily influenced by Adam Neumann's direct engagement with Masayoshi Son, overriding traditional due diligence [5]. Stewart Butterfield's past success with Flickr similarly bolstered Slack's early investments [7].

Geographic Bias and Market Familiarity. Investment paths that included geographic connections showed significant variability ($p < 0.05$), suggesting that location serves as a heuristic for market familiarity and risk assessment.

Examples: SoftBank's early $20M investment in Alibaba capitalized on geographic familiarity with the Chinese market [18], while Tiger Global's early backing of Flipkart reflected strategic market specialization in India [17].

Optimal VC Co-Investment (Saturation Effect). A nuanced, statistically validated effect was observed regarding prior VC involvement. Startups with 1–2 prior VC investors attracted significantly higher path scores ($p < 0.001$), while those with more than three experienced diminishing returns ($p < 0.001$). This indicates an optimal co-investment range, where moderate backing is seen as a quality signal, while excessive investment reduces exclusivity.

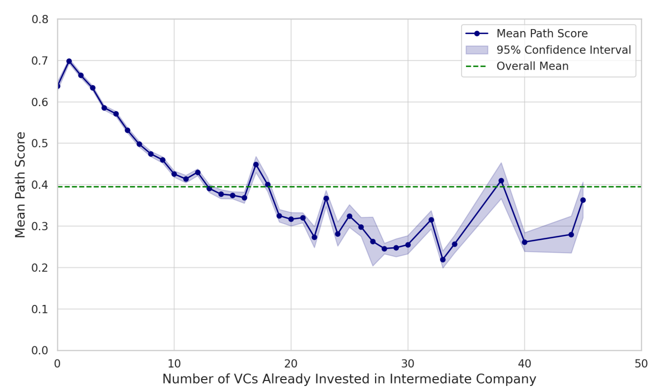

Figure 2: Mean path score by number of prior VC investors, showing optimal attractiveness for 1–2 prior investors.

6 Conclusion

We introduced an explainable recommender system tailored for high-stakes investment decisions, leveraging KPAR's path-based reasoning over a knowledge graph. Our method uniquely provides global interpretability through statistically validated path attention, uncovering key investment signals such as industry familiarity, founder reputation, angel investor influence, optimal VC saturation, and geographic preferences. These insights align with well-documented investment strategies, highlighting the practical relevance of our approach. Notably, our evaluation showed that interpretability does not come at the cost of predictive accuracy, underscoring KPAR's applicability for real-world decision-making in venture investments.

6.1 Future Directions

Future work should explore KPAR's robustness for new investors and startups with limited historical data. Investigating the scalability of KPAR on larger, more complex knowledge graphs is crucial, alongside optimizations for efficient path extraction. Addressing real-time updates to the knowledge graph and their integration into the recommendation process remains an open challenge. Comparative analysis with other KG-based recommenders (e.g., KGAT [20], CKAN [22]) could further substantiate KPAR's competitive strengths in investment scenarios.

Presenter Biography

Prof. Noam Koenigstein is a Professor at the School of Industrial Engineering and Intelligent Systems at Tel Aviv University. His research focuses on deep learning for large-scale recommender systems and explainable AI in high-stakes domains. Over the past year, he has been collaborating with ALPHA10x alongside Veronika Bogina and Amir Wolf, advising on the development of their investment knowledge graph. Prior to his academic role, Prof. Koenigstein led recommendation research teams at Microsoft, where he drove machine learning innovations for Xbox and Microsoft Store. He also served as Senior Vice President and Head of Data Science at Citi Bank's Israeli Innovation Lab, overseeing AI-driven financial technologies.

References

[1] Andreessen Horowitz. 2021. Coinbase Direct Listing. https://a16z.com/coinbase-direct-listing/ Accessed: 2025-05-11.

[2] Giacomo Balloccu, Ludovico Boratto, Christian Cancedda, G. Fenu, and M. Marras. 2023. Faithful Path Language Modeling for Explainable Recommendation over Knowledge Graph. *arXiv preprint arXiv:2310.16452* (2023). https://arxiv.org/abs/2310.16452

[3] Buzzacott. 2017. Four funding lessons from the founding of Facebook. https://www.buzzacott.co.uk/insights/four-funding-lessons-from-the-founding-of-facebook Accessed: 2025-05-11.

[4] Leigh Eytan, Veronika Bogina, Irad Ben-Gal, and Noam Koenigstein. 2024. KPAR: Knowledge-aware Path-based Attentive Recommender with Interpretability. *ACM Transactions on Recommender Systems* (2024).

[5] Fortune. 2023. Beyond WeWork: How A Japanese billionaire VC took $60 billion ... https://fortune.com/2023/11/07/wework-bankruptcy-unicorns-venture-capital-softbank-masayoshi-son-billionaire/ Accessed: 2025-05-11.

[6] Qingyu Guo, Fuzhen Zhuang, Chuan Qin, Hengshu Zhu, Xing Xie, Hui Xiong, and Qing He. 2020. A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering* 34, 8 (2020), 3549–3568.

[7] Inc. 2023. Slack's Stewart Butterfield on Getting to $1 Million in 72 Hours. https://www.inc.com/christine-lagorio/slack-s-stewart-butterfield-on-getting-to-1-million-in-72-hours.html Accessed: 2025-05-11.

[8] Bo Jiang, Junchen Yang, Yanbin Qin, Tian Wang, Muchou Wang, and Weifeng Pan. 2021. A service recommendation algorithm based on knowledge graph and collaborative filtering. *IEEE access* 9 (2021), 50880–50892.

[9] Simerjot Kaur, Ivan Brugere, Andrea Stefanucci, Armineh Nourbakhsh, Sameena Shah, and Manuela Veloso. 2021. Parameterized Explanations for Investor/Company Matching. *arXiv preprint arXiv:2111.01911* (2021).

[10] Johannes Luef, Christian Ohrfandl, Dimitris Sacharidis, and Hannes Werthner. 2020. A recommender system for investing in early-stage enterprises. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*. 1453–1460.

[11] Power in Numbers. 2023. Google's earliest angel investors. https://powerinnumbers.medium.com/googles-earliest-angel-investors-63e2e66f5ccc Accessed: 2025-05-11.

[12] Subarna Shakya Rama Bastola. 2024. Knowledge-Enriched Graph Convolution Network for Hybrid Explainable Recommendation from Review Texts and Reasoning Path. *2024 International Conference on Information and Communication Technology* (2024). https://doi.org/10.1109/ICICT60155.2024.10544384

[13] Reuters. 2018. SoftBank to move ride-hailing stakes worth $20 billion to Vision Fund. https://www.reuters.com/article/us-softbank-group-vision-fund/softbank-to-move-ride-hailing-stakes-worth-20-billion-to-vision-fund-idUSKBN1HW0FG Accessed: 2025-05-11.

[14] Zeinab Shokrzadeh, Mohammad-Reza Feizi-Derakhshi, Mohammad-Ali Balafar, and Jamshid Bagherzadeh Mohasefi. 2024. Knowledge graph-based recommendation system enhanced by neural collaborative filtering and knowledge graph embedding. *Ain Shams Engineering Journal* 15, 1 (2024), 102263.

[15] Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S Yu, and Tianyi Wu. 2011. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *Proceedings of the VLDB Endowment* 4, 11 (2011), 992–1003. https://doi.org/10.14778/3402707.3402736

[16] Zhu Sun, Jie Yang, Jie Zhang, Alessandro Bozzon, Long-Kai Huang, and Chi Xu. 2018. Recurrent knowledge graph embedding for effective recommendation. In *Proceedings of the 12th ACM conference on recommender systems*. 297–305.

[17] The Economic Times. 2023. Tiger trail at Flipkart ends in $3.5 billion profits. https://m.economictimes.com/tech/newsletters/morning-dispatch/tiger-trail-at-flipkart-ends-in-3-5-billion-profits-28-gst-differences-emerge-among-gaming-firms/articleshow/102258524.cms Accessed: 2025-05-11.

[18] The Wall Street Journal. 2014. SoftBank's Alibaba Alchemy: How to Turn $20 Million Into $50 Billion. https://blogs.wsj.com/digits/2014/09/19/softbanks-alibaba-alchemy-how-to-turn-20-million-into-50-billion/ Accessed: 2025-05-11.

[19] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, and M. Guo. 2018. RippleNet: Propagating user preferences on the knowledge graph for recommender systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 1–18.

[20] Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. Kgat: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 950–958.

[21] Xiang Wang, Dingxian Wang, Canran Xu, Xiangnan He, Yixin Cao, and Tat-Seng Chua. 2019. Explainable reasoning over knowledge graphs for recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 5329–5336.

[22] Ze Wang, Guangyan Lin, Huobin Tan, Qinghong Chen, and Xiyang Liu. 2020. *CKAN: Collaborative Knowledge-Aware Attentive Network for Recommender Systems*. Association for Computing Machinery, New York, NY, USA, 219–228.

[23] F. Zhang, N.J. Yuan, D. Lian, X. Xie, and W.-Y. Ma. 2016. Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 353–362.

[24] Xiaoxue Zhao, Weinan Zhang, and Jun Wang. 2015. Risk-hedged venture capital investment recommendation. In *Proceedings of the 9th ACM Conference on Recommender Systems*. 75–82.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009