Lexical Link Analysis and Quantum Intelligence Game (LLAQIG) for Crowdsourcing High-value Information

1. Introduction

Traditionally in network theory, the importance of a node is a form of high-value information. Current methods of ranking high-value nodes require established hyperlinks, citation networks, social networks, or other forms of collective intelligence marked by humans. However, few or no hyperlinks are available for private or proprietary data. Furthermore, high-value information can be different from applications. Current methods mainly score popular information, which are mostly useful for marketing applications. Anomalous information is important for intelligence analysis. When an information provider seeks crowdsourcing from an audience's response to a piece of information, it can be viewed as a strategic cooperation game of two players. The information provider's search for a Nash equilibrium may not achieve a full Pareto efficiency or the so-called optimal social welfare, referred as the Prisoner's dilemma. It is necessary to apply quantum computing and game properties to reach both the Nash equilibrium and optimal social welfare [1].

2. LLAQIG Model

Lexical link analysis (LLA), an unsupervised learning algorithm, discovers word feature pairs, networks, and themes grouped into popular and anomalous ones. Quantum superposition and entanglement of the popular and anomalous themes of LLA determine the value of a piece of information. For two pure strategies for C and D, a superposition would be

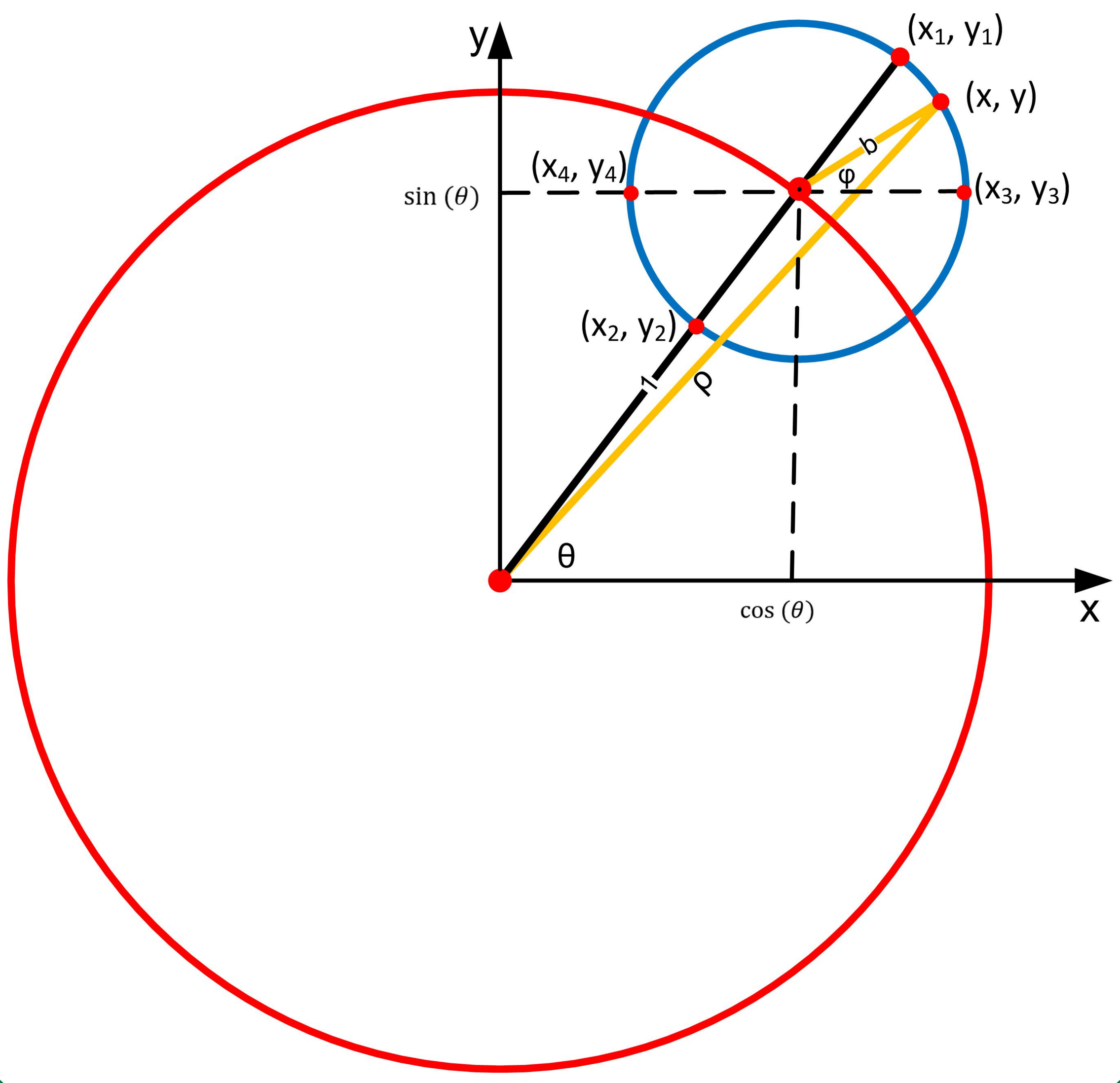
$$|X\rangle = c_0 |C\rangle + c_1 |D\rangle \quad (1)$$

where the coefficients c_1, c_2 are complex numbers.

$$c_0 |C\rangle = e^{i\theta}$$
$$c_1 |D\rangle = be^{i\phi} \quad (2)$$

The magnitude of the superposition as a total reward, i.e., total social welfare ρ, for a quantum game is

$$\rho = \langle X|X\rangle = 1 + b^2 + 2b\cos(\theta - \phi) \quad (3)$$



3. Use Case

Daily business news from 1/1/2021 to 8/31/2021 for about 7000 publicly traded companies (self-player) was used to validate the LLAQIG model. Companies' stock prices are impacted either positively (higher) or negatively (lower) by the business news. Variables include numbers of popular or anomalous word features for each news from LLAQIG , news sentiment, stock's daily return, volume, and other measures.

7. References

[1] Z.W. Sun. The rule for evolution of cooperation in quantum games. *ACTA PHYSICA POLONICA A.*, 116(2), 2009.

5. Comparison with Supervised ML Methods

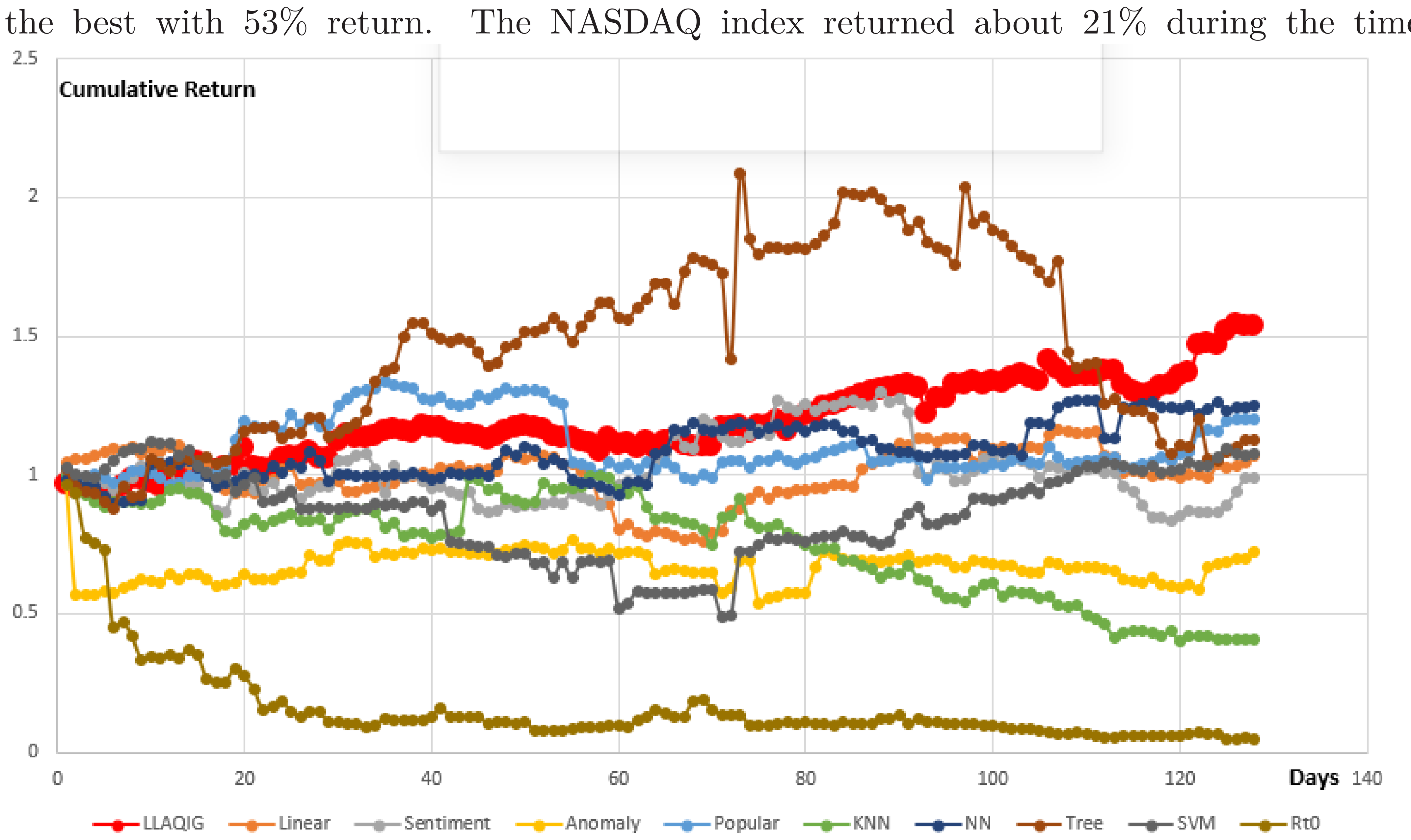
Our LLAQIG is an unsupervised algorithm. Keywords from business news are different among different industries and changing dynamically over time. Traditional supervised machine learning and predictive algorithms are difficult to apply. Nevertheless, for comparison, we first apply our LLAQIG and then compare the results with predictive algorithms from the sciki-learn python package.

Table 2: LLAQIG Results and Comparison

Method	Mean	Variance	CV	CR	Return
LLAQIG	0.0040	0.022	0.18	1.54	54%
NN	0.0021	0.031	0.07	1.25	25%
Popular	0.0018	0.028	0.06	1.20	20%
Tree	0.0026	0.062	0.04	1.12	12%
SVM	0.0022	0.059	0.04	1.08	8%
Sentiment	0.0007	0.039	0.02	0.99	-1%
Anomaly	-0.0007	0.055	-0.01	0.72	-28%
Linear	-0.0010	0.039	-0.03	1.07	7%
Random Forest	-0.0160	0.120	-0.14	0.05	-95%
KNN	-0.0062	0.042	-0.15	0.40	-60%

4. Results

The y-axis shows the cumulative or compound return (CR) for the top stock selected using the unsupervised LLAQIG , i.e., the stock has the highest value from LLAQIG and other scores using comparative supervised algorithms for the 128 days (x-axis). The coefficient variance (CV) is the ratio of mean and variance. CV is a good indication of the performance of stock return. Anomaly and Popular scores alone do not perform as well as LLAQIG . Our LLAQIG performs the best with 53% return. The NASDAQ index returned about 21% during the time period.



6. Conclusions

We demonstrated a quantum machine learning system LLAQIG to discover high-value information from crowdsourcing.

1. Our LLAQIG is unsupervised. LLA can be set up as a game-theoretic framework where a self-player as an information provider against the opponent as the audience responding to the information.
2. We showed quantum entanglement and superposition effect may be in work in real life when an information provider and its audience interact.
3. We validated our LLAQIG using a practical and open source financial data to find the best intrinsic values that reach the Nash equilibrium and maximize social welfare simultaneously.