

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**ScienceDirect**

Procedia Computer Science 162 (2019) 916–923

**Procedia**

Computer Science

[www.elsevier.com/locate/procedia](http://www.elsevier.com/locate/procedia)7th International Conference on Information Technology and Quantitative Management  
(ITQM 2019)

## A risk-aware fuzzy linguistic knowledge-based recommender system for hedge funds

Álvaro Tejada-Lorente<sup>a</sup>, Juan Bernabé-Moreno<sup>a</sup>, Julio Herce-Zelaya<sup>a</sup>, Carlos Porcel<sup>b</sup>,  
Enrique Herrera-Viedma<sup>a,1</sup><sup>a</sup>Andalusian Research Institute in Data Science and Computational Intelligence, University of Granada. E-18071 Granada, Spain<sup>b</sup>Department of Computer Science, University of Jaén, Jaén E-23071, Spain

### Abstract

One of the most difficult tasks for hedge funds investors is selecting a proper fund with just the right level level of risk. Often times, the issue is not only quantifying the hedge fund risk, but also the level the investors consider just right. To support this decision, we propose a novel recommender system, which is aware of the risks associated to different hedge funds, considering multiple factors, such as current yields, historic performance, diversification by industry, etc. Our system captures the preferences of the investors (e.g. industries, desired level of risk) applying fuzzy linguistic modeling and provides personalized recommendations for matching hedge funds. To demonstrate how our approach works, we have first profiled more than 4000 top hedge funds based on their composition and performance and second, created different simulated investment profiles and tested our recommendations with them.

© 2020 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the 7th International Conference on Information Technology and Quantitative Management (ITQM 2019)

**Keywords:** Recommender systems, finance, hedge funds, knowledge-based

### 1. Introduction

Hedge funds are investment mechanisms with a high trading activity and often subject to soft regulation. Their objective is to maximize returns in any possible manner. Often time, hedge funds are based on tax haven countries and willing to use any possibility the market could offer, with no restrictions. In addition, their operations are not exactly risk-averse: massive leverage, simultaneous long and short positions, usage of complex investment instruments, such as derivative securities, options, etc. In their early days, there were no transparency on how hedge funds were operated, but after several scandals [1] some regulations were introduced. At the same time, while hedge funds were originally created for institutional investors and wealthy individuals, the increasing amount of information required and published by regulators, increased the transparency and brought hedge funds closer to wider audiences, especially those willing to take higher risks for higher profits. At present, the amount and variety of information available about hedge funds is such, that investors end up facing the well known information overload problem [2] at the time to decide on personal investments.

---

\*Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .

E-mail address: [viedma@decsai.ugr.es](mailto:viedma@decsai.ugr.es).

Obtaining potentially high profit ratios require a considerable amount of time and expertise and decisions might come with no guarantee for success. On one side, we encounter traditional investments funds which are collective investing entities operating through different financial instruments where users delegate their responsibilities to management companies, normally banks or other specialized entities. They provide professional investment managers, who attempt to deliver sustainable returns with a more conservative risk management approach. On the other side, hedge funds also manage investments on behalf of the investors, but pursuing a higher risk while seeking for potentially bigger returns. Hedge funds usually require a higher minimum initial investment than investment funds. The selection of the appropriate hedge fund to manage our investments is a fairly complicated decision that could be assisted by the appropriate information selection system, which in finance usually translates in the adoption of new technologies [3]. Recommender Systems (RSs) have been proven successful when used on information overload problems, not only in finance but also on similar areas providing personalized solution to users. Some examples of successful adoption can be found in e-commerce [4], health [5], education [6] and others [7, 8].

In this paper we present a fuzzy linguistic knowledge-based recommender system for hedge fund recommendations based on users risk profiles and their preferred industries. We model hedge fund risk based on their results, the industry representation share ratio and their diversification. We assumed an adequate investing experience and knowledge from users, considering that the definition of own hedge funds is reserved for advanced investors. The major contributions and novelties proposed in our solution include:

- The usage of fuzzy linguistic modeling to assist users to specify their preferences by expressing their risk profiles and preferred industries.
- The hedge fund risk assessment, minimizing the required data, using current and past performance, share ratios and diversification indicators per industry.
- The knowledge-based recommendation strategy based on risk assessments on hedge funds and its adaptability to the user's needs.

The paper is organized as follows. The background on hedge funds, fuzzy linguistic modeling and recommender systems is presented in Section 2. Section 3 describes the system and its functioning. In section 4 we present different simulated scenarios and the discussion of the results. Finally, we present our conclusions and point to further research lines based on our proposed method.

## 2. Background

In the background section we provide the required reference information to describe our system. First, we will introduce recommender systems and a brief categorization, then we will explain the main characteristic from hedge funds followed by a description of the fuzzy linguistic modeling.

### 2.1. Recommender systems

RSs aid users in the items selection process by filtering non desired information based on users or item profiles [9]. Many successful examples in several fields have proven them as useful tools: e-commerce [4, 10], health [5, 11], learning [6, 12], etc. RSs based their functioning on building profiles with information they have from users and information they have about items to recommend. Such information can be obtained on different ways depending on the nature of the system, either *explicitly* by users inputs or admin inputs or *implicitly* by the normal functioning of the system, logs, rates or external information that could be used [13].

RSs can be categorized based on the recommendations generation approach they follow. In [14], Burke proposed to categorize them based on 4 different recommendation techniques: *Collaborative*, that is, the system uses the rating similarities provided by users. *Content-based*, where the system generate the recommendations using the features related with items and the previous rates from users. *Demographic*, the recommendations are generated based on the demographic characteristic from users. Finally, in the approach that we will use in our system, *Knowledge-based*, the recommendations are generated based on user's preferences having additional knowledge that could assist on the matching of this preferences either on the user side or the item side.

## 2.2. Hedge funds

Hedge funds are private investment business which operate managing the investment of other parties. The main characteristic of hedge funds is the exclusivity on the investors selection. Unlike mutual funds, hedge funds are much more flexible in the types of securities hold and the type of position taken, they can invest in all the traded derivative securities, they may hold long and short positions, invest on equities or debt, and they assume more risk with less diversified investments as well as bigger leverage [15]. They need to be registered with the Commodity Futures Trading Commission (CFTC) by membership in the National Futures Association [16]. In this setup, hedge funds have the authority to act on behalf of customers on futures, securities accounts and options.

Hedge funds have the obligation to fill on quarterly basis the so called 13-F forms. These forms contain all equity assets under management of at least a value of \$100 million. Authorities warn about the drawbacks of relying just on these reports since hedge funds are not obliged to report about their short positions and other specific type of investments<sup>1</sup>.

## 2.3. Fuzzy linguistic modeling

When the information is provided as a linguistic form, normally is not expressed in a quantitative mode but in a qualitative manner. For that, in [17], Zadeh introduced the concept of *linguistic variable* which has been successfully used for modeling qualitative information [18, 19, 20].

The problem of information loss associated with the fuzzy linguistic modeling [17] has been addressed in [21], by the 2-tuple approach for fuzzy linguistic modeling, by representing through continuous model the data that allows to reduce the information loss.

This model defines the functions  $\Delta(\beta)$  and  $\Delta^{-1}(s_i, \alpha)$  to transform between numeric values ( $\beta$ ) and 2-tuples  $((s_i, \alpha))$ , as well as a set of negation, comparison and aggregation operators [21].

The differences between uncertainty degrees when using linguistic modeling arise the problem of different cardinalities on the linguistic term sets [22]. For that, a multi-granular 2-tuple fuzzy linguistic modeling based on the concept of linguistic hierarchy is proposed in [22].

A *Linguistic Hierarchy*,  $LH$ , is a set of levels  $l(t, n(t))$ , where each level  $t$  is a linguistic term set with disparate granularity  $n(t)$  from the other levels. The order follows is related with the granularity of the levels. A level can be constructed from its predecessor as:  $l(t, n(t)) \rightarrow l(t + 1, 2 \cdot n(t) - 1)$ . The set of transformation functions between labels from different levels is bijective in order to avoid the information loss[22]

In [22] authors remark that the family of transformation functions between labels from different levels is bijective, guaranteeing that the transformations between levels are produced without loss of information in a linguistic hierarchy.

## 3. Proposal description

The problem we are dealing with has very specific idiosyncrasy. On the one hand, recommendations about investments in hedge funds are not required to be received on a high frequency base but more on demand. When a potential investor is willing to take a decision and fills or updates its profile based on their current knowledge or needs, the system will estimate the best possible recommendations for the investor and provide them to him. On the other hand, any time new reports about hedge funds are introduced in the system, the item profiles need to be newly calculated to include the newer information about performance and risks.

In this section we are going to describe our risk aware hedge fund fuzzy linguistic knowledge-based recommender system. We can distinguish three key elements: hedge fund risk modeling by industry and performance, the fuzzy linguistic profiles representation and the knowledge-based recommender scheme.

<sup>1</sup><https://www.sec.gov/answers/about-lawshtml.html#secexact1934>

### 3.1. Hedge fund risk aware model

This module identifies the industries importance and performance, as well as the diversification rate per hedge fund.

The hedge funds data used for risk modeling were extracted from the public quarterly 13F-HR forms, filed by institutional management investments that must be reported to the United States Securities and Exchange Commission (SEC) regulations<sup>2</sup>.

In order to estimate the risk profile per hedge fund, we have analyzed the industry sector proportion share per hedge fund, as well as the performance in time for each one.

We have the 13F-HR reports of a total of 4878 hedge funds, we have categorized them into a total of 60 different industries plus an ‘Other’ category where is compound all investment entities with no clear industry.

Let  $HF = \{HF_0, \dots, HF_N\}$  be the set of  $N$  different hedge funds,  $I = \{I_0, \dots, I_M\}$  be the set of  $M$  different industries and  $E = \{E_0, \dots, E_P\}$  be the set of  $P$  different investments entities. Each investment entity will belong to at least one or more industries. Each hedge fund will have their investments on one or more investments entities,  $HF_i \rightarrow \{E_{i0}, \dots, E_{in}\}$ . In order to simplify it, we will group the investments entities per industries per hedge fund, being  $HF_i = \{I_j, \dots, I_k\}$  the vector that will represent hedge fund  $i$ , where the industry group  $j$  in hedge fund  $i$  will contain certain group of investment entities,  $I_{ij} = \{E_x, \dots, E_y\}$ .

In finance, the risk use to be measured by the standard deviation in the performance through time. A higher standard deviation is considered to be exposed to a higher risk. We would like to go a bit further, since we are building a recommender system and we look for recommendations which provide a good performance, we have built a specific attribution model to distribute the risk per industry based on all hedge funds composition and their diversification, as well as to consider the performance and risk through time per hedge fund.

So we define the Risk Index (RI) from a hedge fund  $i$  as follows:

$$RI_i = \alpha_1(\lambda_1 Performance_i * \lambda_2 Risk_i) * \alpha_2 \Sigma[(RiskIndustryExposure_{ij})]$$

$Performance_i$  is the performance from hedge fund  $i$  expressed as:  $Performance_i = \gamma_1 PPM_{5Y} * \gamma_2 PPM_Y * \gamma_3 PPM_Q$ . Being  $\gamma$  the parameter to balance the importance of performance in time and  $PPM_x$  the Performance Per Month for a certain time, in our case, a quarter, a year and 5 years.  $Risk_i$  is the risk from hedge fund  $i$  expressed as:  $\delta_1 SD_{5Y} * \delta_2 SD_Y * \delta_3 SD_Q$ . Being  $\delta$  the parameter to balance the importance of risk in time, measured as  $SD_x$ , that is, standard deviation for a certain time, and  $\lambda$  the parameter used to module the importance of performance over risk or viceversa.

Note that, since we will need to build it as a vector per hedge fund, we will express it per industry without the sum as:

$$RI_{ij} = \alpha_1(\lambda_1 Performance_i * \lambda_2 Risk_i) * \alpha_2 (RiskIndustryExposure_{ij})$$

$RiskIndustryExposure_{ij}$  represent the exposure of hedge fund  $i$  to the industry  $j$  regarding the presence inside hedge fund and the proportional weight of industry  $j$  among all the others hedge funds. We consider that a higher exposure risk comes when a industry is over represented regarding the same industry among all hedge funds and when the diversity of different investment entities in the industry is smaller than usual.

For that, we define it as:

$$RiskIndustryExposure_{ij} = \beta_1((PV_{ij} - \mu_{PV_j})/\sigma_{PV_j}) * \beta_2((PNE_{ij} - \mu_{PNE_j})/\sigma_{PNE_j})$$

Where we try to measure the statistical dispersion through the distance from the mean values measured in standard deviations, that is, their confidence interval. Being,  $PV_{ij}$  the percentage value of the industry  $j$  from the total value in hedge fund  $i$ , and  $\sigma_{PV_j}, \mu_{PV_j}$  the standard deviation of the percentage values of the industry  $j$  among all hedge funds and their average. Meanwhile,  $PNE_{ij}$  is representing the percentage of number of investment entities in the industry  $j$  from the total value in hedge fund  $i$ , and  $\sigma_{PNE_j}, \mu_{PNE_j}$  the standard deviation of the percentage of number of investment entities of the industry  $j$  among all hedge funds and their average.

<sup>2</sup><https://www.sec.gov/cgi-bin/browse-edgar?action=getcurrent&datea=&dateb=&company=&type=13F-HR&SIC=&State=&Country=&CIK=&owner=include&accno=&start=200&count=100>

### 3.2. Fuzzy linguistic profiles representation

In order to represent the information in our recommender system, we have built fuzzy linguistic profiles for users preferences and for items. So, for that we will select two different label sets from a linguistic hierarchy of two levels. The first level will have 5 labels to represent the risk degree ( $S_1 = S^5$ ) and the second level will have 7 labels to represent the importance degree ( $S_2 = S^9$ ). The concepts represented are the following:

- **Importance degree** of a industry regarding the final investment, which is assessed in  $S^9$ .
- **Risk degree** that the user is willing to take regarding hedge fund selection, which is assessed in  $S^5$ .
- **Similarity degree** between resources and users, which is assessed in  $S^9$ .

The linguistic hierarchy proposed has the following linguistic terms sets:

- $S^5 = \{b_0 = \text{None} = N, b_1 = \text{Low} = L, b_2 = \text{Medium} = ME, b_3 = \text{High} = H, b_4 = \text{Maximum} = M\}$ .
- $S^9 = \{c_0 = \text{None} = N, c_1 = \text{VeryLow} = VL, c_2 = \text{MoreLessLow} = MLL, c_3 = \text{Low} = L, c_4 = \text{Medium} = ME, c_5 = \text{High} = H, c_6 = \text{MoreLessHigh} = MLH, c_7 = \text{VeryHigh} = VH, c_8 = \text{Maximum} = M\}$ .

The item representation will be carried out by a classification of hedge funds per industries (61 industries). The system will obtain an internal representation based on the risk aware model explained above per industry and hedge fund, that is, the *Risk Index* per industry. Thus, to represent an item  $i$ , that is, a hedge fund  $i$ , we use a vector model  $VR_i = (VR_{i1}, VR_{i2}, \dots, VR_{i61})$ , where  $VR_{ij} \in S^9$  will show the *Importance degree* of an Industry inside each hedge fund. Note that, for the item profile representation we do not use the aggregated  $RI_i$  but the individual  $RI_{ij}$  per industry.

In order to get user's preferences we will ask them to select explicitly 5 industries where they believe hedge funds should be mainly focused on, based on their knowledge or preferences ( $S^9$ ). We will also require them to provide the general level of risk they are willing to take with their investments ( $S^5$ ). After that, we will adjust the risk level expressed by the user from  $S^5$  level to  $S^9$  level, and we will aggregate it to all the industries on the user vector with the same relation expressed by  $\alpha_1$  and  $\alpha_2$ . This will result on the user vector  $VU_i = (VU_{i1}, VU_{i2}, \dots, VU_{i61})$ .

### 3.3. Knowledge-based recommender scheme

Based on the risk-aware approach we have followed that has provided us with highly valuable extra insights about hedge funds behavior and composition, the recommendations generation will be created by a knowledge-based recommender system.

We have proceeded to match the item and users vector profiles through similarity measures, choosing as the more suitable the Cosine similarity measure [23] over the Pearson [24] similarity measure, since the profiles consist on a vector of features where just a small set will make the bigger contribution. Considering that we are working with linguistic labels, we use the standard cousin measure, but adapted to a linguistic context:

$$\sigma_l(V_1, V_2) = \Delta(g \times \frac{\sum_{k=1}^n (\tilde{h}_1 \times \tilde{h}_2)}{\sqrt{\sum_{k=1}^n (\tilde{h}_1)^2} \times \sqrt{\sum_{k=1}^n (\tilde{h}_2)^2}})$$

where  $V_1$  and  $V_2$  are the two user or profile vector to compare,  $\sigma_l(V_1, V_2) \in S_1 \times [-0.5, 0.5]$ ,  $g$  is the granularity of the term set used to express the Importance degree ( $S_2$ ),  $n$  the number of industries,  $\tilde{h}_i = \Delta^{-1}(v_{ik}, \alpha_{vik})$  y  $(v_{ik}, \alpha_{vik})$  is the 2-tuple linguistic value of the term  $k$  in the vectors  $V_i$ , with  $i = 1, 2$ .

#### 3.3.1. Knowledge-based recommendations

There are two scenarios when the recommendations are generated:

1. Any time a new set of reports with new information is made available by the authorities, usually every quarter. Risk Indexes need to be estimated again for each eligible hedge fund  $i$ .
2. Any time a user  $e$ , new or already existing, create or modify their preferences.

When any of the both scenarios take place, we estimate  $\sigma_1(VR_i, VU_e) \in S_2$  for each user. As  $S_2 = S^9$ , we consider that  $i$  is related with  $e$  if  $\sigma_1(VR_i, VU_e) > (s_5^9, 0)$ .

After all the similarity degrees are estimated, the top 5 hedge funds on a descending order by similarity degree, included on the related set of hedge fund per user, will be presented to every user as eligible. Coupled with each hedge fund recommended, a visual report will be presented displaying their Risk Indexes as well as their performance and industries distribution.

#### 4. System testing and discussion

In order to test our system we have simulated 3 different user investment profiles. The users will be called ‘user 1’, ‘user 2’ and ‘user 3’ and they will have *low*, *medium* and *maximum* as risk profiles. We have selected 5 specific industries per profile with specific interests as can be seen in table 1.

| User 1               |  |             |
|----------------------|--|-------------|
| Investment           |  | \$100.000   |
| Risk Profile         |  | Low         |
| Industry Speculation | Business Service   | High        |
|                      | Chemical and allied products                               | Medium      |
|                      | Insurance carriers   | Low         |
|                      | Communication  | High        |
|                      | Depository institutions                                    | Medium      |
| User 2               |  |             |
| Investment           |  | \$500.000   |
| Risk Profile         |  | Medium      |
| Industry Speculation | Industrial and commercial machinery and computer equipment | High        |
|                      | Transportation equipment                                   | Medium      |
|                      | Holding and other investment offices                       | Low         |
|                      | Communication  | Very High   |
|                      | Petroleum refining and related industries                  | Medium      |
| User 3               |  |             |
| Investment           |  | \$1.000.000 |
| Risk Profile         |  | Maximum     |
| Industry Speculation | Others   | Very High   |
|                      | Oil and gas extraction                                     | Medium      |
|                      | Insurance carriers   | Low         |
|                      | Educational services                                       | Low         |
|                      | Petroleum refining and related industries                  | High        |

Table 1. User Profiles 1,2 and 3

To obtain the recommendations we have calibrated our system with data until beginning of 2019. The parameters for the Risk Index estimation have been adjusted as follows:  $\alpha_1 = 0.5, \alpha_2 = 0.5, \lambda_1 = 0.4, \lambda_2 = 0.6, \gamma_1 = 0.5, \gamma_2 = 0.3, \gamma_3 = 0.2, \delta_1 = 0.5, \delta_2 = 0.3, \delta_3 = 0.2, \beta_1 = 0.7, \beta_2 = 0.3$ .

Having the system with the users and items profiles loaded as well as the required parameters we have generated the top 5 recommendations per users.

In order to validate the results of the recommendations, we have checked any possible investment on each of the top 5 hedge fund recommendation per user with real data from the last quarter results per hedge fund available, that is, data from 30/06/2019.

Below we can see the top 5 hedge funds recommended per user and their performances on the mentioned quarter (expressed on percentage over the investment):

**User 1: \$100.000**

- 1 Close Asset Management Ltd -0.73
- 2 Fernwood Investment Management, LLC 2.74
- 4 Mraz, Amerine & Associates, Inc. 5.09
- 3 S. MUOIO & CO. LLC 2.01
- 5 Bedel Financial Consulting, Inc. 5.85

**User 2: \$500.000**

- 1 GAMCO INVESTORS, INC. ET AL 1.30
- 2 Clarius Group, LLC 7.67
- 3 GABELLI & Co INVESTMENT ADVISERS, INC. -0.23
- 4 AR ASSET MANAGEMENT INC 3.71
- 5 Hillsdale Investment Management Inc. 1.85

**User 3: \$1.000.000**

|   |                                  |      |
|---|----------------------------------|------|
| 1 | Mraz, Amerine & Associates, Inc. | 5.09 |
| 2 | Sunbelt Securities, Inc.         | 5.15 |
| 3 | EVERENCE CAPITAL MANAGEMENT INC  | 6.21 |
| 4 | Swiss National Bank              | 5.33 |
| 5 | CHICAGO EQUITY PARTNERS LLC      | 3.80 |

As we can see, for *User 1*, 4 of 5 hedge funds have reported benefits, with an average of 2.99%. *User 2* as well is having 4 of 5 hedge funds reporting benefits, one of them with a high 7.67% performance in just one quarter. Finally, *User 3*, which was the user with the maximum risk profile, obtained an average of 5.11% performance on his/her recommendations.

Furthermore, in Fig. 1 we present the performance of the top 15 hedge funds related to user 1, 2 and 3. For that, we have combined them all making a total of 39 hedge funds that were eligible for being recommended to some of the three users.

|                    |                                  |                                 |                                  |                                   |                              |                                  |                      |                           |
|--------------------|----------------------------------|---------------------------------|----------------------------------|-----------------------------------|------------------------------|----------------------------------|----------------------|---------------------------|
| USCA RIA LLC       | DekaBank Deutsche Girozentrale   | TD ASSET MANAGEMENT INC         | Mraz, Amerine & Associates, Inc. | PARSONS CAPITAL MANAGEM... INC/RI | PETTYJOHN, WOOD & WHITE, INC | COMERICA BANK                    |                      |                           |
|                    | Congress Park Capital LLC        | Capital Wealth Planning, LLC    | MONTAG & CALDWELL, LLC           | LCNB CORP                         | CHICAGO EQUITY PARTNERS LLC  | Executive Wealth Management, LLC | GABELLI FUNDS LLC    |                           |
| Clarius Group, LLC |                                  | EVERENCE CAPITAL MANAGEMENT INC | Swiss National Bank              | Qube Research & Technologies Ltd  | AR ASSET MANAGEM... INC      | Yorktown Manage... & Researc...  | Pendal Group Ltd     | WALTER & KEEN... FINAN... |
|                    | Bedel Financial Consulting, Inc. | Sunbelt Securities, Inc.        | BUCKHEAD CAPITAL MANAGEM... LLC  | Fernwood Investment Manageme...   | ENVESTNET ASSET MANAGEM...   | S. MUO... & CO...                | Hillsdale Investm... | F3Logic, LLC              |
|                    |                                  |                                 |                                  |                                   | WEAL... ENH... ADVI...       | Candri... Luxem...               | W... IN... Y... A... | GAMC... LET... T...       |

Fig. 1. Performance per hedge funds related with users 1, 2 and 3

It is worth to mention that the results are only for the first quarter of the year and due to the underlying nature of hedge funds the are not definitive but a good indicative of the tendency.

**5. Conclusions**

This study presents a fuzzy linguistic knowledge-based recommender system, which makes use of public data to create risk-aware hedge funds profiles and recommends them to users based on their investment preferences. The data set used has been built from the public data extracted from officials 13F-HR reports submitted to the SEC and after being analyzed in detail. A *Risk Index* per hedge fund has been defined and estimated to include the concept of performance into hedge fund risk profile. It includes the exposure per industry as well as its diversification compared to the strategies followed by all other hedge funds. The system captures the investments preferences from users regarding their preferred risk profile, as well as the industries they would like to have as main part on their investment and recommends the 5 hedge funds that are closer to the given set of preferences.

The main advantage of the proposal is the fact that the risk evaluation is made by the system before hand and then considered in the recommendation process. The *Risk Index* reflects two main characteristics, the past performance together with the risk exposure based on results and per industries. We have simulated three different

profiles, with different industries preferences and different risk profiles (low, medium and maximum). For that, we had set up our system with data until beginning of 2019 and tested the recommendations obtained from the system on the last quarter reports regarding their performance. We have seen that more than 86% of hedge funds recommended had performed positively. Additionally, they have obtained a positive 3.65% performance as average. The maximum performance obtained on the reviewed quarter have been 7.67% by *Clarius Group, LLC* and the minimum of -0.73% by *Close Asset Management Ltd.*

As future work, we consider addressing the risk individually per investment equity that belongs to each industry as well as obtaining further data sources to increase the reliability in addition to the 13F report.

## Acknowledgments

This paper has been developed with the FEDER financing under Project TIN2016-75850-R. The Titan Xp used for this research was donated by the NVIDIA Corporation.

## References

- [1] W. Fung, D. A. Hsieh, A primer on hedge funds, *Journal of empirical finance* 6 (3) (1999) 309–331.
- [2] A. Edmunds, A. Morris, The problem of information overload in business organizations: a review of the literature, *International Journal of Information Management* 20 (2000) 17–28.
- [3] Á. García-Crespo, J. L. López-Cuadrado, I. González-Carrasco, R. Colomo-Palacios, B. Ruiz-Mezcua, Sinvlio: Using semantics and fuzzy logic to provide individual investment portfolio recommendations, *Knowledge-Based Systems* 27 (2012) 103–118.
- [4] Y. Zhang, H. Abbas, Y. Sun, Smart e-commerce integration with recommender systems, *Electronic Markets* 29 (2) (2019) 219–220.
- [5] S. Hors-Fraile, O. Rivera-Romero, F. Schneider, L. Fernandez-Luque, F. Luna-Perejon, A. Civit-Balcells, H. de Vries, Analyzing recommender systems for health promotion using a multidisciplinary taxonomy: A scoping review, *International journal of medical informatics* 114 (2018) 143–155.
- [6] Á. Tejeda-Lorente, J. Bernabé-Moreno, C. Porcel, P. Galindo-Moreno, E. Herrera-Viedma, A dynamic recommender system as reinforcement for personalized education by a fuzzy linguistic web system, *Procedia Computer Science* 55 (2015) 1143–1150.
- [7] F. Z. Benkaddour, N. Taghezout, F. Z. Kaddour-Ahmed, I.-A. Hammadi, An adapted approach for user profiling in a recommendation system: Application to industrial diagnosis., *IJIMAI* 5 (3) (2018) 118–130.
- [8] E. R. Núñez-Valdez, D. Quintana, R. G. Crespo, P. Isasi, E. Herrera-Viedma, A recommender system based on implicit feedback for selective dissemination of ebooks, *Information Sciences* 467 (2018) 87–98.
- [9] R. Burke, Hybrid web recommender systems, P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.): *The Adaptive Web*, LNCS 4321 (2007) 377–408.
- [10] S. S. Li, E. Karahanna, Online recommendation systems in a b2c e-commerce context: a review and future directions, *Journal of the Association for Information Systems* 16 (2) (2015) 72.
- [11] B. Esteban, A. Tejeda-Lorente, C. Porcel, M. Arroyo, E. Herrera-Viedma, Tplufib-web: A fuzzy linguistic web system to help in the treatment of low back pain problems, *Knowledge-Based Systems* 67 (0) (2014) 429 – 438.
- [12] M. Hsu, A personalized english learning recommender system for esl students, *Expert Systems with Applications* 34 (2008) 377–408.
- [13] N. Good, J. Schafer, J. Konstan, A. Borchers, B. Sarwar, J. Herlocker, J. Riedl, Combining collaborative filtering with personal agents for better recommendations, in: *Proceedings of the Sixteenth National Conference on Artificial Intelligence*, 1999, pp. 439–446.
- [14] R. Burke, Hybrid recommender systems: Survey and experiments, *User Modeling and User-Adapted Interaction* 12 (2002) 331–370.
- [15] S. J. Brown, W. N. Goetzmann, R. G. Ibbotson, Offshore hedge funds: Survival and performance 1989-1995, Tech. rep., National Bureau of Economic Research (1997).
- [16] W. Fung, D. A. Hsieh, Empirical characteristics of dynamic trading strategies: The case of hedge funds, *The Review of Financial Studies* 10 (2) (1997) 275–302.
- [17] L. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning. Part I, *Information Sciences* 8 (1975) 199-249, Part II, *Information Sciences* 8 (1975) 301-357, Part III, *Information Sciences* 9 (1975) 43-80 (1975).
- [18] F. J. Cabrerizo, J. A. Morente-Molinera, W. Pedrycz, A. Taghavi, E. Herrera-Viedma, Granulating linguistic information in decision making under consensus and consistency, *Expert Systems with Applications* 99 (2018) 83–92.
- [19] F. Cabrerizo, R. Al-Hmouz, A. Morfeq, A. Balamash, M. Martnez, H.-V. E., Soft consensus measures in group decision making using unbalanced fuzzy linguistic information, *Soft Computing* 21:11 (2017) 3037–3050.
- [20] S. Massanet, J. V. Riera, J. Torrens, E. Herrera-Viedma, A new linguistic computational model based on discrete fuzzy numbers for computing with words, *Information Sciences* 258 (2014) 277–290.
- [21] F. Herrera, L. Martínez, A 2-tuple fuzzy linguistic representation model for computing with words, *IEEE Transactions on Fuzzy Systems* 8(6) (2000) 746–752.
- [22] F. Herrera, L. Martínez, A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multi-expert decision-making, *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics* 31(2) (2001) 227–234.
- [23] J. Ye, Cosine similarity measures for intuitionistic fuzzy sets and their applications, *Mathematical and computer modelling* 53 (1-2) (2011) 91–97.
- [24] L. Sheugh, S. H. Alizadeh, A note on pearson correlation coefficient as a metric of similarity in recommender system, in: *2015 AI & Robotics (IRANOPEN)*, IEEE, 2015, pp. 1–6.