Personalized Finance Advisory through Case-based Recommender Systems and Diversification Strategies

Cataldo Musto[a,*], Giovanni Semeraro[a], Pasquale Lops[a], Marco de Gemmis[a], Georgios Lekkas[b]

[a] *Universita degli Studi di Bari "Aldo Moro", Department of Computer Science*
[b] *Objectway Financial Software, Italy*

Abstract

Recommendation of financial investment strategies is a complex and knowledge-intensive task. Typically, *financial advisors* have to discuss at length with their wealthy clients and have to sift through several *investment proposals* before finding one able to completely meet investors' needs and constraints. As a consequence, a recent trend in wealth management is to improve the advisory process by exploiting recommendation technologies. This paper proposes a framework for recommendation of asset allocation strategies which combines *case-based reasoning* with a novel diversification strategy to support financial advisors in the task of proposing diverse and personalized investment portfolios. The performance of the framework has been evaluated by means of an experimental session conducted against 1172 real users, and results show that the yield obtained by recommended portfolios overcomes that of portfolios proposed by human advisors in most experimental settings while meeting the preferred risk profile. Furthermore, our diversification strategy shows promising results in terms of both diversity and average yield.

Keywords: Recommender Systems, Case-based Reasoning, Personalization,

*Corresponding author
Email addresses: cataldo.musto@uniba.it (Cataldo Musto), giovanni.semeraro@uniba.it (Giovanni Semeraro), pasquale.lops@uniba.it (Pasquale Lops), marco.degemmis@uniba.it (Marco de Gemmis), lekkas@objectway.it (Georgios Lekkas)

DOI: 10.1016/j.dss.2015.06.001

1. Background and Motivations

Financial services firms such as banks, brokerages, family offices, life insurance companies and trusts offer investment services to their clients and help them reach their objectives. Such investment services typically include advisory on investment strategies, discretionary portfolio management in which clients delegate portfolio management to experts, sales of financial products offered by the firm or third parties, collection and transmission of trading orders to financial markets. Clients are classified in segments based on their available assets as Retail, Affluent, High Net Worth Individuals (HNWI), or Ultra High Net Worth (UHNW) individuals and are treated differently, with different products and services proposed to meet their needs.

After the 2008 financial crisis, all financial services firms increased their focus on investment services, as they are profitable but do not involve credit risk nor stress banks' capital requirements. At about the same time specific regulations such as MiFID[1] in Europe or Retail Distributions Review (RDR)[2] in the UK, were established to protect investors and their assets. Firms wanting to expand their market share and meet regulatory requirements had to invest heavily in new processes and IT platforms to improve their offerings, quality of service and compliance. Indeed, to know the clients and to deliver them personalized investment proposals is today considered as an essential facet of a fruitful and effective advisory strategy [2]. IT investments were oriented towards increasing transparency, delivering better and more timely client reporting, but did not influence the investment decision-making process.

In the last few years the rapidly moving scenario has been further revolutionized by the technology trends subsumed under the term Digitization, which

[1] http://en.wikipedia.org/wiki/Markets_in_Financial_Instruments_Directive
[2] http://www.fsa.gov.uk/rdr

despite hesitations, will deeply and unavoidably transform the wealth management industry [4]. The effects of digitization include reduction of the number of physical branches and the transition of business transactions to online channels. A part of the digitization trend are online advice sites, sometimes called ”robo-advisors”, that let clients get advice online, anytime at a lower cost [1]. Online advice platforms support a Do-It-Yourself (DIY) attitude of clients and put pressure on professional advisors who follow the traditional wealth management model of personal interactions and paper-based processes [3].

To cope with online competition and with pressure on costs coming from the increased regulatory requirements, advisors should now make the most of their time and maximize the quality of their advice while operating with efficiency. Efficiency is particularly important when working with clients of the Affluent segment, who are much more numerous than the HNWI and UHNW segments. An example of platform-supported efficiency are advisors receiving intelligent help to quickly sift through past data and exploiting the past experience of the firm to give the best possible solutions to their clients. This brings in the idea that recommendation technologies could be adapted in the investment services context and be the advisors' assistant in the new operating environment.

2. Goal and Contributions

As proved by many success stories, Recommender Systems (RS) [5] can provide users with high-quality personalized suggestions and can effectively support people in real-time decision making tasks. However, the application of such technology in the *financial domain* is neither trivial neither straightforward, since some peculiarities of this domain make hard to put into practice the most common recommendation paradigms such as the content-based (CB) [6] and the collaborative filtering (CF) [7] ones.

Indeed, in this particular setting each user can be just modeled through his *risk profile*[3] along with some demographical features, while each financial prod-

[3] The Risk Profile is defined as ”an evaluation of an individual or organization's willingness

uct is described through a *rating*[4] provided by credit rating agencies, an average *yield* on different time intervals and the *category* it belongs to. This makes a pure CB strategy very likely to fail, since content-based information is too poor and not meaningful to feed a CB recommendation algorithm. Moreover, the over-specialization problem [8], typical of CB recommenders, may collide with the fact that turbulence and fluctuations in financial markets suggest to change and diversify the investments over time. Similarly, CF algorithms can hardly be adopted because of the well-known *sparsity* problem, which arise when is very difficult to identify the neighbors of the target user.

However, the main reason that makes CB and CF strategies very likely to fail lies in the absence of a real user history (in terms of positive and negative ratings) for the financial domain. Indeed, each user typically keeps its asset allocation strategy constant for a long period of time, so it is not possible to accumulate enough ratings to trigger a classical recommendation process relying on the analysis of previous preferences of the users or on the analysis of ratings patterns within the community of users. Due to these dynamics, it is necessary to focus on different recommendation paradigms. Knowledge-based Recommender Systems (KBRS) [9], for example, provide users with recommendations by typically matching preferences and domain constraints with a set of possible solutions. This insight fits well with the financial domain since there is a clear relationship between the risk profile of the target user and the asset classes he is more inclined to invest in [10].

However, due to the complexity of the *knowledge acquisition* step, which is mandatory for KBRS, the research in the area shifted the focus to a subclass of KBRS called case-based recommender systems (CBRS) [12]. CBRS avoid the bottleneck of explicit knowledge acquisition by adopting *case-based reasoning* (CBR) [11], a problem solving methodology that tries to solve new problems

---

to take risks". Typically, this value is obtained by conducting the above mentioned standard MiFiD questionnaire.

[4] http://en.wikipedia.org/wiki/Credit_rating

by re-using specific past experiences stored in some example cases. Specifically, CBRS recommendations rely on the retrieval and adaptation of the suggestions proposed in similar settings, which are drawn from a set - called *case base* - of (effective) previously proposed solutions.

This paper proposes a framework for *recommendation of asset allocation strategies* relying on case-based reasoning. The framework is the outcome of a joint research with Objectway Financial Software aimed at improving the advisory process implemented in OFS ADVICE[5], a platform for investor-centric wealth management. OFS ADVICE defines and tailors an investment proposal in terms of *asset allocation* and *product recommendations* that meet all the investor's objectives.

The proposed framework merges the advantages of KBRS with the simplicity of a recommendation process which avoids explicit knowledge acquisition. Furthermore, a strategy to provide users with diverse investment solutions is integrated, in order to effectively deal with market fluctuations and flocking. In the experimental session our framework has been compared to a k-NN baseline as well as to recommendations provided by human advisors in both in-vitro and in-vivo ex-post evaluation.

To sum up, the contributions of the paper can be summarized as follows:

1. It introduces a novel framework for recommendation of asset allocation strategies;
2. It evaluates the effectiveness of CBRS recommendation strategies in a special (and, to best of our knowledge, not yet evaluated) domain;
3. It proposes a greedy diversification algorithm able to diversifying the investment strategies over time;
4. It evaluates the effectiveness of the framework through an extensive *ex-post evaluation*;

[5] http://www.objectway.com/EN/financial-software/FS-Advisors-Network-MIFID-advice.asp

The paper is organized as follows: Section 3 provides an overview of the literature. The framework for recommendation of asset allocation strategies is described in Section 4, while Section 5 provides a thorough description of the experimental design as well as the outcomes of the evaluation. Finally, the conclusion and the research to be carried on are sketched in Section 6.

3. Related Work

Recommendation of financial investment strategies is a very controversial and complex topic. Generally speaking, this research line has a strong relationship with the area of *human decision-making* [13]. It is not by chance that many researchers tried to investigate human behavioral patterns in the areas of both financial decision-making [17, 18, 20] and asset allocation [19]. The complexity of the task is also confirmed by several work that aim to learn whether some relationship exists between psychological traits of wealthy clients and the investment proposals they chose [15]. Given that some research already underlined the (positive) role of case-based reasoning strategies in human-decision making tasks [14], CBR was chosen as backbone of our framework for financial recommendation. However, the adoption of this strategy in the financial domain has been poorly investigated, with the exception of the model proposed by Chuang [21], which exploits CBR for bankruptcy detection.

The first attempts towards the usage of CBR in recommendation-related tasks date back to the early 2000s in the e-commerce [38], restaurant [36], and tourism [37] domains. In the first case CBR is adopted to support users' choices through a conversational interface, while in the others CBR is triggered according to users' preferences, typically expressed as a logical query on the case base. The only difference between our approach and state of the art ones lies in the way user preferences are represented. In our setting, user is represented according to her financial-based as well as demographical charachteristics, while in the above mentioned attempts user is modeled through her preferences in the food domain or through her travel wishes (town, hotel features, weekdays, activities and so on). Furthermore, differently from our framework, none of the state of the art approaches takes into account diversity issues.

As regards recommender systems in the financial domain, in [16] Yu proposes an architecture of a decision-support system for the financial domain. This is a very preliminary attempt, since no technical and methodological details are provided for the implementation of each recommendation step. The main contribution in the area is due to Felfernig et al. [33], who proposed a framework for the development of KBRS which is the main building block of FSAdvisor [32], a platform for financial services recommendation. The most distinguishing aspect of our work lies in that, differently from [33], we adopted CBRS instead of KBRS as recommendation paradigm, thus avoiding explicit and hand-crafted knowledge acquisition.

Some recent work focusing on the development of frameworks [30] as well as recommendation models [29] based on CBR confirmed the interest of the research community in this area as well as the effectiveness of such approach in several scenarios.

Most of the literature tackles the portfolio recommendation problem by adopting Artificial Intelligence techniques. In [34], Gonzalez-Carrasco et al. adopted fuzzy logic to automatically classify the *risk profile* of the user according to social and psychological facets. Similarly, in [35] the authors used fuzzy logic and association rule mining to generate a portfolio based on the analysis of stock market trends. Fuzzy set theory is also applied in [24] to build a financial portfolio. Other attemps are based on genetic algorithms [22, 23], neural networks [25, 26] and multi-criteria decision making [27, 28]. Moreover, a recent work by Taghavi et al. [31] proposed an agent-based framework to provide financial recommendations relying on a hybrid technique which combines macro-economic factors and techniques for trend predictions with a classical collaborative filtering algorithm.

Differently from most of these approaches, our framework is more oriented to financial advisors since our goal is not to automatically build a portfolio, but rather to help advisors in filtering the proposals on the basis of previous

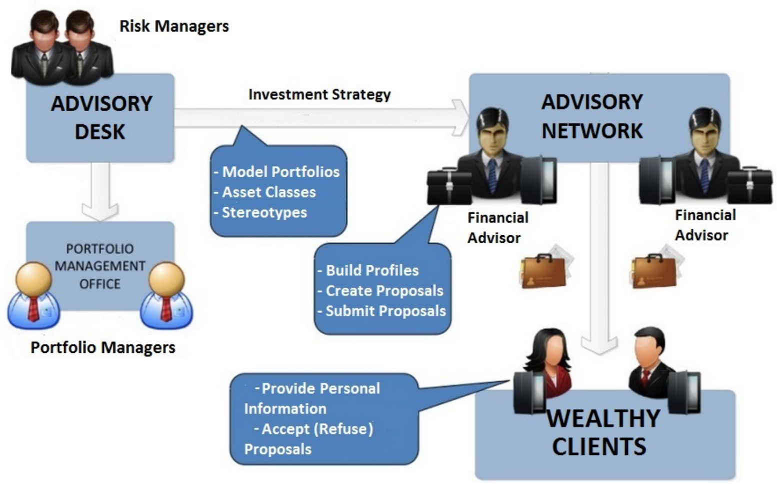


Figure 1: A classical workflow for financial recommendations

cases, in order to identify the most suitable (and diversified) available portfolios. Moreover, we focused on the task of recommending asset allocation strategies instead of recommending particular financial products since, as proved by previous research, 50% of investment returns are explained by asset allocation [39]. Finally, differently from our work, none of the above mentioned approaches neither compared the performance of the recommended portfolio to that of portfolios recommended by human advisors nor performed a real *ex-post* evaluation with real users.

4. Methodology

In a classical workflow not fostered by a recommender system (Figure 1), a central *advisory desk* operated by investment strategists estimates market trends and future returns of each asset class on the basis of mathematical models. Clients are grouped in *stereotypes* on the basis of some attributes (e.g. risk attitude, investment expertise, investment horizon). A set of fixed basic proposals, called *model portfolios*, is handed over to advisors, together with a static mapping of client stereotypes to model portfolios.

Financial advisor cannot change the fundamentals of this process, which ensures regulatory compliance and suitability of proposals to clients. But advisors still face the problem of selecting one of the model portfolios variants and adapting it to each particular clients' constraints, needs and desires (for example, a client may dislike investing in a particular geography or industry). Typically, the final portfolio proposal consists of a set of asset classes with their percentages: an example of the output of the advisory process is provided in Table 1. The process may conclude here, but more often it proceeds to transform asset class recommendations in portfolios of financial products for each specific asset class. So a key task of the advisor is to ensure the correct mapping of clients to stereotypes and adapt the generic model portfolio proposal by further discussing with the client. Some advisors have experience of decades and have a long history of consistent returns and high client satisfaction. They may not fully agree with the outlook of the advisory desk; they apply their own tactics in the context of the broader strategy, changing percentages of asset classes in the final solution.

Asset Class	**Percentage**
Euro Bond	40%
High Yield Bond	30%
Euro Stocks	15%
Emerging Market Stocks	15%

Table 1: An example of investment portfolio

Given this scenario, our recommendation framework is based on the insight that a financial advisor, before arriving to a final proposal, could benefit from the analysis of portfolios of clients similar to the current client, proposed in the past by herself or by other advisors, accepted by clients and with a good track record. This insight perfectly fits with the principles of case-based reasoning, since it tries to solve new problems by re-using specific past experiences [42]. In our specific case, case-based reasoning is exploited to drive the recommendation process on the basis of a *case base* of previously proposed investments.

A case-based recommendation workflow is typically structured in five different steps: RETRIEVE, REUSE, REVISE, REVIEW and RETAIN. In the RETRIEVE step a set of problems already solved in the past, sharing common charachteristics with the new one, is drawn from the case base. Next, in the REUSE step, the solutions previously adopted for such problems are extracted. The REVISE and REVIEW steps aim at adapting those solutions to fit the specific constraints of the new problem. Finally, in the RETAIN step, the solution obtained by mining previously solved problems is proposed to the user. According to user feedback, good solutions are stored in the case base in order to exploit them again in future to solve similar problems.

4.1. Case-based Reasoning for Recommendation of Asset Allocation Strategies

Our recommendation process, sketched in Figure 2, is based on an adaptation to the financial domain of the above described workflow.

Formally, given a case library C, each case c_i is a triple $\langle u_i, p_i, f_i \rangle$, where u_i is a representation of a user, p_i is a representation of the portfolio, and f_i is a feedback assessment. Each user u_i is represented as a vector of eight features: *risk profile*, inferred through the standard MiFiD questionnaire, *investment goals*, *investment horizon*, *investment experience*, *financial assets*, *sex*, *advice type* and *age*. A thorough description of the features is provided in Section 5. The first five features are represented on a five-point ordinal scale, from *very low* to *very high*, while *sex* and *advice type* are represented in a binary fashion (normal advice = 0, extended advice = 1, female = 0, male = 1). Finally, age is represented as a numerical variable. Each portfolio p_i is represented as the distribution of the asset classes that compose it, such as *Euro Bond*, *High Yield Bond*, *Emerging Markets Stock Options* and so on, along with their percentage. Feedback f_i is the yield obtained by the portfolio.

Given such a representation and given a *new problem*, that is to say, a new user requesting personalized financial advice, our recommendation pipeline is structured as follows:

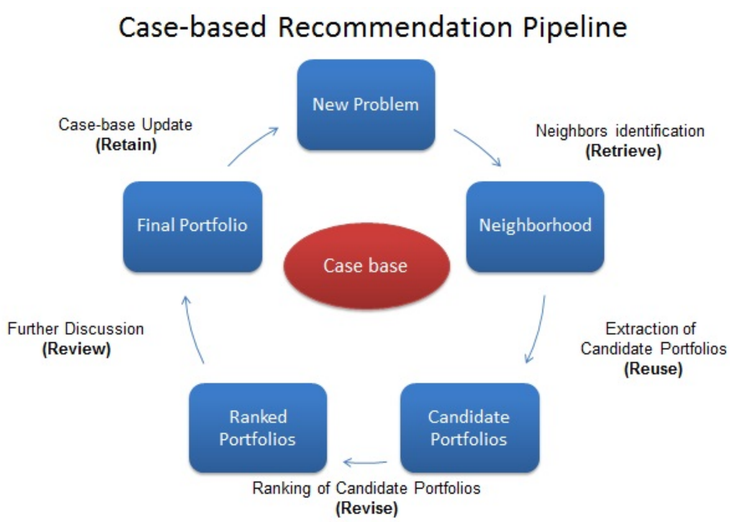


Figure 2: Our Case-based Recommendation Pipeline

(1) Retrieve and Reuse: the goal of the RETRIEVE step is to identify similar (already solved) cases which can be potentially useful to tailor the investment proposal. This task can be split in two parts: first, users with similar needs (called *neighbors*) are retrieved, then solutions proposed to the neighbors are extracted.

Generally speaking, the first part can be tackled as a classical *retrieval task*, that is to say, given a vector-space representation of the target user built upon the weight of each feature (very low = 1, very high = 5 for ordinal features, 0 or 1 for binary ones), the process returns a set of k similar users (called *neighbors*) $N = \{n_1 \dots n_k\}$.

For this step, two different approaches have been implemented: *user match* and *cosine similarity*. By following the first strategy, all the cases whose users exactly share the same features are retrieved. In alternative, *cosine similarity* [40] is exploited to retrieve the most similar users.

Let $\vec{u}$ be a vector-space representation of the target user and $\vec{c}$ be a vector-space representation of another user already stored in the case base, *cosine similarity* can be computed as follows:

$$\cos(\vec{\mathbf{u}}, \vec{\mathbf{c}}) = \frac{\sum_{i=1}^{n} \vec{\mathbf{u}}_i \vec{\mathbf{c}}_i}{\sqrt{\sum_{i=1}^{n} (\vec{\mathbf{u}}_i)^2} \sqrt{\sum_{i=1}^{n} (\vec{\mathbf{c}}_i)^2}} \tag{1}$$

Users are ranked according to their descending $\cos(\vec{\mathbf{u}}, \vec{\mathbf{c}})$ scores, and the first k neighbors are labeled as *neighbors*.

In the REUSE step, let N be the set of the k neighbors returned by the RETRIEVAL module, the portfolio p_{n_i} of each neighbor n_i, $1 \leq i \leq k$, is gathered and labeled as *candidate solution*.

(2) Revise: the candidate solutions retrieved by the first step are typically too many to be effectively consulted by a human advisor. Thus, the goal of the REVISE step is to further refine the set of *candidate solutions* in order to obtain the final set of *portfolios* which are proposed to the target user. Typically, this is done by adopting clustering, ranking or filtering strategies on the set of candidate solutions. In this work we propose several revise strategies to rank the list of the portfolios, which are discussed in the next section.

(3) Review and Retain: in the REVIEW step human advisor and investor can further discuss and modify the portfolio in order to come to the *final solution*. Once the final recommended portfolio has been agreed, according to some heuristics (e.g., whether the generated yield is over a certain threshold) the solution can be stored in the case base and used as input to solve similar cases in the future.

4.2. Revise Strategies for Recommendation of Asset Allocation Strategies

The proposed framework implements five different revise techniques.

(a) Basic Ranking: portfolios are ranked according to the output produced by the RETRIEVE step. If *user match* is adopted as similarity measure, portofolios are ranked on the basis of the agreement date (from the newest to the oldest), otherwise they are ranked according to their *cosine similarity* scores (Formula 1). The first k portfolios are returned to the advisor as *final solutions*.

(b) Clustering: retrieved portfolios are clustered according to the k-means clustering algorithm [41]. The centroids of the k clusters are returned as *final* *solutions*. This technique is supposed to provide the advisor with more diverse solutions, since similar portfolios are clustered together.

(c) Greedy Diversification: this strategy implements the diversification algorithm described in [43]. The approach tries to diversify the *final solutions* by iteratively picking from the original set of *candidate solutions* the ones with the best trade-off between similarity and diversity, as it is our interest to discover and promote investment solutions that are qualitatively different, possibly result of particular tactic insights. Let u be the target user, F be the set of final solutions, C_{retr} be a set of previously retrieved cases, at each step the algorithm ranks the retrieved neighbors by calculating the *quality* score of each case $c_i = \langle u_i, p_i, f_i \rangle \in C_{retr}$, as follows:

$$Quality(u, c_i, F) = cos(\vec{u}, \vec{u_i}) * relDiv(p_i, F) \tag{2}$$

$$relDiv(p_i, F) = \sum_{j=1}^{|F|} \frac{1 - cos(\vec{s_i}, \vec{f_j})}{|C|} \tag{3}$$

where $cos(\vec{u}, \vec{u_i})$ is the cosine similarity between users and $relDiv(p_i, F)$ is the average diversity of the portfolio p_i with respect to the previously picked solutions. At each step, the solution with the highest *quality score* is removed from the set of candidate solutions and is stored in F. Since at the first iteration F is empty, $relDiv(n, F) = 1$ for all the neighbors. Thus, the first pick is the item with the highest similarity. Next, at each iteration, the solution with the best score is chosen.

(d) FCV: this strategy adapts the *Interest Confidence Value* (ICV) proposed in [44] to the financial domain. As ICV, originally adopted for the restaurant domain, calculates how close are the attributes of the restaurant to those the user already expressed interest in the past (according also to a drift factor, which emulates the process of people losing interest over time), our *Financial Confidence Value (FCV)* calculates how close to the optimal one is the distribution of the asset classes in a portofolio, according to the average historical yield obtained by each class. Given a set of asset classes A, for each portfolio

p, the set P of the asset classes which compose it and its complement $\overline{P}$, are computed. Next, FCV is formally defined as:

$$FCV(p) = Y(p)^{log(\lambda)+1} \tag{4}$$

$$Y(p) = \sum_{i=1}^{|P|} p_{a_i} * y_{a_i} \qquad \lambda = \frac{\sum_{i=1}^{|P|} y_{a_i}}{\sum_{k=1}^{|\overline{P}|} y_{a_k}} \tag{5}$$

where p_{a_i} and y_{a_i} are the percentage and the average yield of the *i-th* asset class in the portfolio, respectively. $Y(p)$ is the total yield obtained by the portfolio, and λ is a drift factor which calculates the ratio in terms of average yield between the asset classes in the portfolio and those which are not in. For values of $\lambda \geq 1$, it acts as a boosting factor (for $\lambda \ll 1$, it acts as a dumping factor). Through this strategy, all the *candidate solutions* are ranked according to the FCV score and the *Top-k* solutions are returned to the advisor.

(e) FCV + Greedy: this combined strategy first uses the greedy algorithm to diversify the solutions, then exploits FCV to rank the portfolios and obtain the *final solutions*.

In the experimental session the effectiveness of all revise strategies has been evaluated.

5. Experimental Evaluation

An extensive series of experiments has been carried out to validate the performance of our framework. The experiments had a threefold goal:

1. Analysis of the influence of each parameter of the framework (*similarity measure*, *feature combinations*, *revise and diversification strategies*) on the performance of recommended portfolios;
2. Comparison of the performance of recommended portfolios to that of portfolios proposed by a human advisor;

3. *Ex-post* comparison of the best-performing configuration to the portfolios proposed by a human advisor after three and after six months from the agreeement date.

5.1. Dataset and Experimental Design

Experiments were performed by exploiting a dataset of 1172 real (anonymous) users, who chose portfolios with financial advisors between June 2011 and June 2013. The dataset has been made available by Objectway Financial Software and is publicly available for download[6]. Each case in the *case base* is represented by adopting the formalism previously introduced in Section 4. Each user is modeled through the features reported in Table 2. Each portfolio consists of 19 different asset classes along with their percentage. The yield of a portfolio of asset classes is the weighted average of the yield of each asset class in the portfolio. The yield of an asset class is measured by the performance of the benchmark linked to and representative of that asset class. This is in accordance to industry standard practice where the performance of a policy portfolio, composed of asset classes is the weighted performance of the passive returns of each asset class [48]. Feedback assessments are calculated on the basis of the average yield generated by each portfolio from the agreement date to January 2014.

To provide users with recommendations, a *leave-one-out* design has been adopted, that is to say, at each run the case base has been built by exploiting all the agreed portfolios with the exception of the one agreed by the target user. Cosine similarity and user match are used as similarity measures for the RETRIEVAL step, while all the previously described ranking strategies are implemented in the REVISE step. Furthermore, in the experimental evaluation different combinations of features have been evaluated. For the sake of simplicity, we define three families of features: BASIC (features 1-5, including all financial-based features), EXTENDED (features 1-6, it adds the advice type to the

[6]http://www.di.uniba.it/~swap/financialrs_data_uniba.zip

basic set) and COMPLETE (features 1-8, it adds the demographic information to the *extended* set). As regards Risk-control, it is already encoded in the retrieval step of the entire process, thus we can avoid to adopt as evaluation measure a risk-adjusted performance (such as the Sharpe Ratio) since we are already sure that the recommendation will be compliant to user risk profile.

#id	**feature**	**type**	**domain**
1	risk profile	ordinal	$[very\ low, low, normal, high, very\ high]$
2	investment goals	ordinal	$[very\ low, low, normal, high, very\ high]$
3	investment horizon	ordinal	$[very\ early, early, normal, long, very\ long]$
4	investment experience	ordinal	$[very\ low, low, normal, high, very\ high]$
5	financial assets	ordinal	$[very\ low, low, normal, high, very\ high]$
6	advice type	nominal	$[normal, extended]$
7	sex	nominal	$\{male, female\}$
8	age	integer	$[18 \ldots 80]$

Table 2: Description of the features adopted to represent users

Statistical differences have been assessed by adopting a *paired t-test* on the average monthly yield of each portfolio (calculated as previously described), with $p < 0.05$.

5.2. Experiment 1: influence of parameters

The performance of recommended portfolios has been compared on the basis of different parameters of the model: *feature combinations*, *similarity measure*, *neighboorhod size* and *revise strategie*.

5.2.1. Ranking Effectiveness

First, we evaluated the **ranking effectiveness** of our approach. Given that our recommender system is supposed to help financial advisors in filtering the available portfolios and providing the wealthy client with the best available proposal, the ability of the algorithm to rank the best portfolio in the first

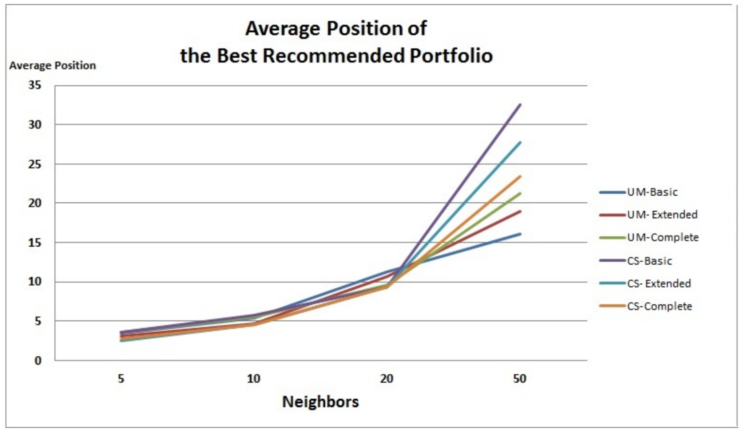
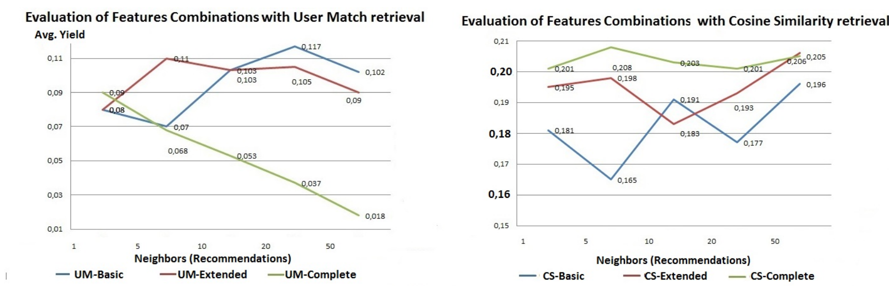


Figure 3: Average Position of the Best Recommended Portfolio

positions is a very important feature. Figure 3 shows the average position of the best recommended portfolio out of a list of n investment proposals (n = *5, 10, 20, 50*). In this setting, the *best* portfolio is the one which obtained the *best average yield* from the agreement date. In this experiment, all the above introduced feature combinations are exploited. The prefixes *UM* and *CS* are used when user match and cosine similarity are the adopted retrieval strategy, respectively.

A quick analysis of the results shows that for n=*5*, n=*10* and n=*20* the ranking does not differ in a significant way, since the best portfolio is always around the middle of the recommendation list, regardless of the features combination and the retrieval strategy. However, with n=*50*, the effectiveness of the *cosine similarity* retrieval drops down in favor of *user match.* This result is probably due to the fact that the adoption of a geometrical model introduces some noise in the retrieval process, especially in the tail of the recommendation list, since portfolios with a poor similarity are proposed to the advisor as *candidate recommendations.*

On the other hand, a less flexible approach based on precise matching can avoid this issue since only the portfolios with a perfect matching in terms of features are retrieved. A statistical analysis performed on the results showed



(a) User Match retrieval

(b) Cosine Similarity retrieval

Figure 4: Evaluation of different feature combinations

that up to *n=20* there is no significant gap between user match and cosine similarity. Thus, we can state that both strategies can be effectively adopted in our recommendation scenario, since it is likely that a human advisor is not willing to scroll a list of more than 20 candidate proposals.

5.2.2. Analysis of Retrieval Strategies

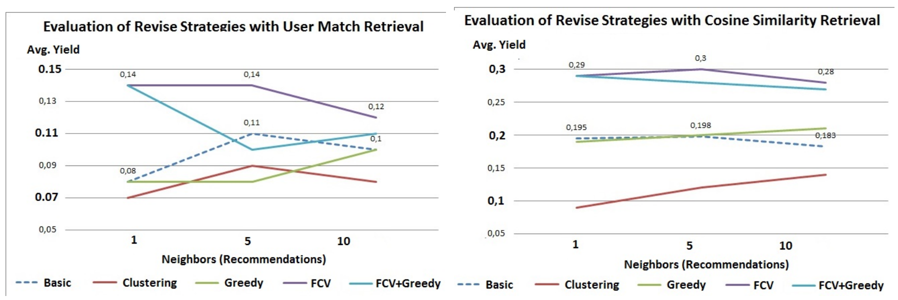
Next, we evaluated how each **retrieval strategy** influences the average yield obtained by the recommended portfolios. Also in this experimental setting, the performance of the portfolios was compared using different *feature combinations*, *similarity measures* and *neighboorhood sizes*. Figure 4a shows the results of the *user match (UM)* retrieval technique with five different neighborhood thresholds and three different combinations of features. The scores reported in the plot represent the average yield obtained by the portfolios agreed by the first n retrieved neighbors (n = *1, 5, 10, 20, 50*).

By analzying the results, it emerges that the introduction of demographical features negatively influences the effectiveness of the recommendation model, especially when the neighbors size increases. Indeed, with *n=10*, *n=20* and *n=50* the yield obtained by the UM-Complete configuration is significantly lower than both UM-Basic and UM-Extended. On the other side, results show that the introduction of a feature based on the kind of advice (UM-Basic vs. UM-Extended) does not influence the overall yield, with the exception of *n=5*. By the way, regardless the significance of the gaps, it emerges that the UM-Basic configuration with *n=20* gets the best results, even if a statistical analysis shows that the neighborhood size does not affect the overall yield.

As regards the adoption of a geometrical retrieval strategy, Figure 4b shows that, unlike the previous experiment, the use of a richer representation can improve the yield obtained by recommended portfolios. Specifically, CS-Complete configuration significantly outperforms CS-Extended with *n=10*, *n=20* and *n=50* and CS-Basic for all neighborhood sizes. These results suggest that a more flexible retrieval strategy based on a vector-space representation can effectively model the information coming from all the available facets of the users, since by exploiting a larger and richer set of features our recommender system is able to retrieve better cases which can provide the final user with a higher yield. Also for CS-based retrieval the statistical analysis confirmed that the neighborhood size does not affect the overall yield. This means that through our approach the advisors do not need to consult a long recommendations list, since even with only *n=1* or *n=5* it is possible to get good suggestions.

Finally, by comparing the best performing configuration obtained by *user match* and *cosine similarity* strategies (CS-Complete vs UM-Basic), it emerges that the adoption of a geometrical retrieval strategy can *significantly* outperform user matching in terms of average yield in all the configurations, since the gap between the strategies is always around 0.1% (on a monthly basis) in favor of CS (0.20% yield for CS vs 0.11% yield for UM, in their best-perfoming configuration).

Even if this experimental result is partially contradictory with the outcomes provided by the analysis of the ranking effectiveness, this can be justified by the fact that the adoption of a geometrical model makes the retrieval process more flexible, thus many relevant portfolios can be introduced in the recommendation list even when the features describing the case do not perfectly match the target scenario. This probably moves down the best available proposal in the recommendation list, but makes the whole list more precise and useful for the financial advisor, since it contains all the proposals able to provide the user with



(a) User Match retrieval (b) Cosine Similarity retrieval

Figure 5: Evaluation of different revise strategies

the best average yield.

5.2.3. *Analysis of Revise Strategies*

Next, we compared the effectiveness of the **revise strategies**. Given that the *neighborhood size* did not significantly affect the overall yield of the portfolios in none of the previous experiments, only the performance of the best-perfoming configurations, CS-Complete and UM-Basic, with smaller neighborhoods have been reported. This choice is also justified by the fact that the recommendation algorithm is supposed to support the financial advisor, thus it makes sense to provide her with only a small list of possible proposals. Results of the comparison are reported in Figure 5a, which shows the average yield obtained by each re-ranking strategy by picking the first n portfolios (n=1, 5, 10) from a neighborhood of fixed size (50) through the user match retrieval and the UM-Basic combination of features.

The main outcome of this experiment is that our novel FCV score outperforms the basic user match retrieval for all neighborhood sizes. Results are particularly significant for *n=1* (with an increase of 75%, from 0.08 to 0.14,) and for *n=5* (with an increase of 27%, from 0.11 to 0.14). Furthermore, the gap between the configurations is statistically significant ($p < 0.05$) for both *n=1* and *n=5*. Our FCV is also the best-perfoming configuration for *n=10*, even if with a smaller (and not significant) improvement. Given that the largest improvement has been obtained with smaller neighborhood sizes, this experiment confirms the effectiveness of the approach, since our novel FCV strategy showed to be really able to put up in the recommendation list the portfolios containing the most promising asset classes. On the other side, the adoption of both diversification and clustering strategies, aimed at providing the advisor with more diverse investment solutions, did not bring any benefit. This result can be due to the fact that these strategies cluster many good recommendations based on similar (good) asset classes, and propose diversified (but worse) solutions to the advisor. Conversely, the insight of combining FCV with the Greedy Diversification strategy got promising results: even if the absolute yield is worse for *n=5* and *n=10*, no statistical differences emerged in any configuration. In the next experiment this outcome will be further investigated in order to analyze the relationship between obtained yield and standard diversity measures.

Similarly, the same experimental protocol has been applied to the best-perfoming configuration based on cosine similarity (CS-Complete). The results of the experiments are plotted in Figure 5b, which confirms most of the outcomes already discussed for UM retrieval. Indeed, also in this experiment the ranking strategy based on FCV obtained the best results for all neighborhood sizes, even with a greater (and statistically significant, $p < 0.05$) improvement with respect to the baseline (around 50% improvement for *n=1*, *n=5* and *n=10*, from 0.19 to 0.29 in terms of average monthly yield). As for the previous experiment, the adoption of clustering did not bring any benefit and the use of the combined FCV + Greedy strategy got promising results. Differently from UM, the integration of a Greedy strategy to diversify CS-based results provides a comparable recommendation accuracy, since the gap is not significant for any neighborhood size. This result is probably due to the fact that CS retrieval produces a more promising list of candidate solutions than UM. In turn, this influences the Greedy strategy which is able to select good and diversified investment solutions. This aspect will be further discussed in the next experiment, as well.

A final overview of the outcomes is shown in Figure 6: it clearly emerges

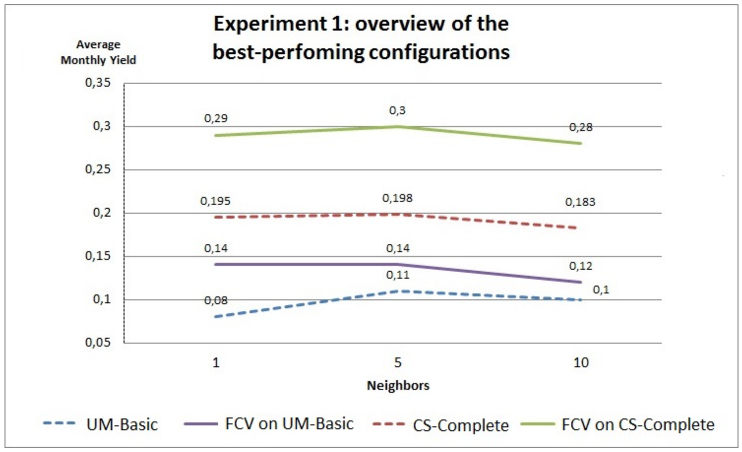


Figure 6: Overview of the best-perfoming configurations of Experiment 1

that the application of FCV can significantly outperform the overall yield of the portfolios generated by both baselines (represented by dashed lines). By the way, regardless the relative improvement, the plot shows that a retrieval approach based on cosine similarity provides the best recommendations, since the application of FCV on the portfolios retrieved through a geometrical model can lead the average yield up to 0.3% on a monthly basis.

To sum up, it is possible to state that the best-performing configuration is the one based on cosine similarity and FCV ranking strategy, with a richer case representation based on all the available features.

5.2.4. Analysis of Diversification strategies

Next, we analyzed the ability of each retrieval and revise strategies to provide the financial advisor with a set of candidate solutions as more **diversified** as possible. In literature several metrics to measure the diversity of a recommendation algorithm have been presented [45]. Among the available ones, we chose the Intra-List Diversity (ILD) [46]. Differently from other well-known diversity metrics, such as the Aggregate Diversity [47], the goal of the ILD is to evaluate how different from each other the items in the recommendation list are. This insight perfectly fits with the financial scenario, since our goal was to

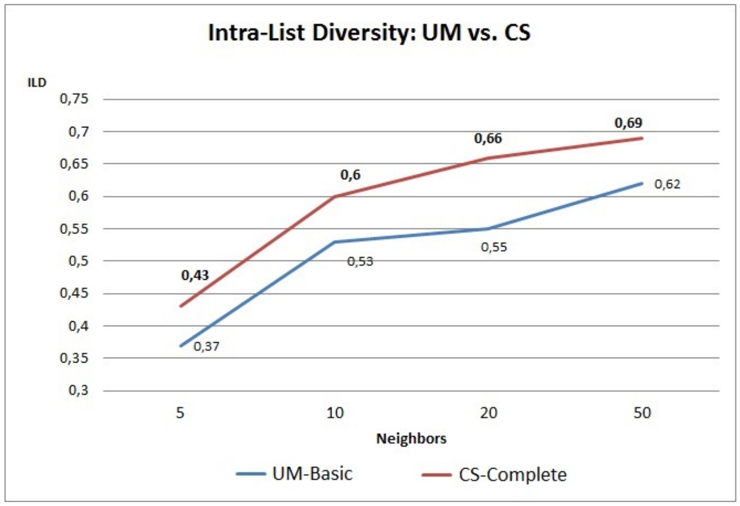


Figure 7: Comparison of UM and CS in terms of ILD

measure to which extent each algorithm is able to provide wealthy clients with diverse investment solutions. Formally, let R be a list of recommendations, and $sim(r_i, r_j)$ be a similarity score between items r_i and r_j both in R, ILD is calculated as follows:

$$ILD(R) = 1 - \frac{\sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} sim(r_i, r_j)}{2n} \tag{6}$$

Figure 7 reports the ILD scores for both UM and CS retrieval strategies. with four different neighborhood sizes (n= *5, 10, 20, 50*). The first outcome of the experiment is that a clear relationship emerges between ILD and the size of the neighborhood, regardless the strategy adopted to retrieve the portfolios. This result was somehow expected since the probability that solutions with higher diversity are included in the list increaseas as the number of recommendations grows. A statistical test showed that the improvement in terms of ILD is significant ($p < 0.05$) when the size of the recommendation list grows from n=*5* to n=*10*.

As regards the retrieval strategy, it emerged that the adoption of CS retrieval results leads to a statistically significant improvement of approximately 10%

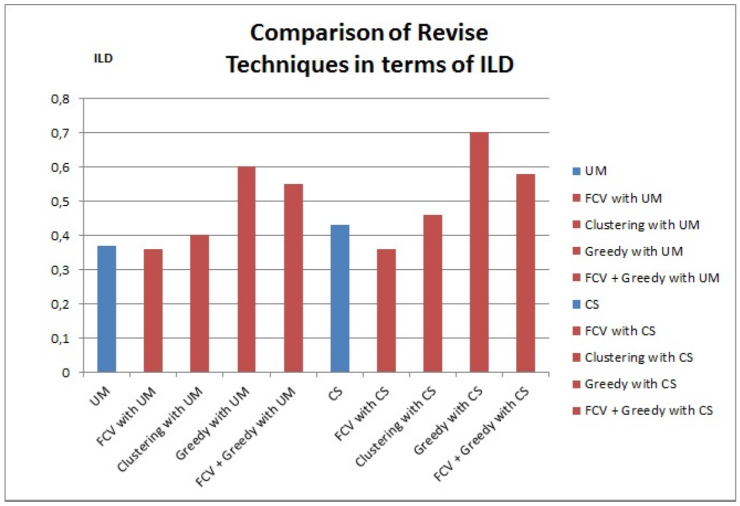


Figure 8: Comparison of revise strategies in terms of ILD

with respect to UM retrieval. Also this result was expected, since a geometrical model provides a more flexible retrieval mechanism which can easily introduce diverse (but sometimes noisy) items in the recommendation list. The results reported in Figure 6 have been exploited as baselines to evaluate the ability of each REVISE strategy to diversify the recommendation list. Specifically, we compared the ILD of recommendation lists of $n=5$ built through all the available re-ranking strategies from a fixed neighborhood of size 50. Results are provided in Figure 8.

The main outcome of the experiment is that the classical Greedy diversification is actually able to largely improve the diversity of the recommendation lists (around 70% improvement on UM retrieval and around 50% improvement on CS retrieval), even if this often leads to a significant worsening of the yield of the portfolios, as already discussed. Also in this experiment our clustering algorithm does not provide benefit, since the little improvement in terms of ILD is not balanced by the overall yield, which gets significantly worse.

As regards FCV, results show that the adoption of a revise strategy based on the analysis of the optimal asset classes produces a loss in diversity. This was an expected outcome, since the rationale behind FCV is to re-rank the rec-

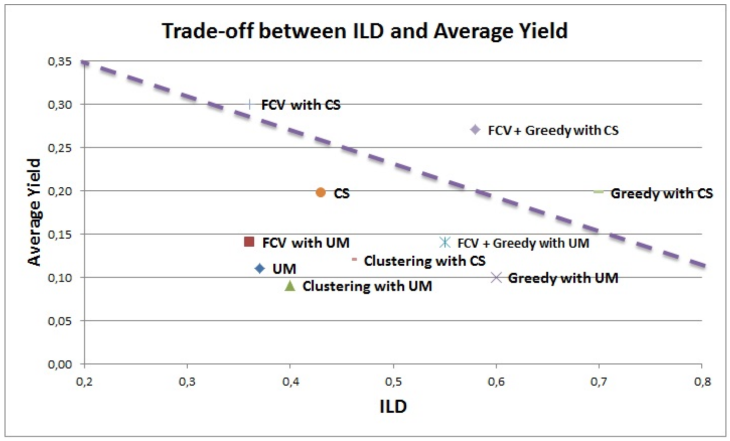


Figure 9: Trade-off between ILD and Average Yield

ommendation list by preferring portfolios containing the asset classes with the best performance in the past. Clearly, this leads to a list of portfolios which are more likely to be very similar to each other. Finally, interesting outcomes emerged by analzying the ILD obtained by the FCV+Greedy strategy, since the idea of combining FCV with a diversification algorithm produces a very diverse (a significant improvement with respect to the baselines was noted for both UM ans CS retrieval) and effective recommendation list (no significant worsening in terms of overall yield, as discussed in the previous experiment). This final outcome further confirms the goodness of such a combined strategy in providing the financial advisor with good and diverse recommendations. Finally, Figure 9 shows the trade-off between ILD and Average Yield for all the configurations. The best-performing settings are those over the frontier (represented by a dashed line). It clearly emerges that the FCV-based strategy is the one able to maximize the average yield, while the simple Greedy strategy provides the best diversity (but sacrificing the performance of the portfolios). As expected, the FCV+Greedy strategy gets the best compromise between both aspects, since it can provide users with good and diverse portfolio recommendations.

To sum up, it is possible to state that the configuration which provides the best diversity is the Greedy one. Experimental results showed that FCV + Greedy *strategy is able to lead to diversified recommendations which can provide the user with good average yield as well.*

5.3. Experiment 2: comparison to baselines

The results of Experiment 1 represent a good picture of the overall effectiveness of the recommendation framework in terms of both diversity of the suggestions and overall yield generated by recommended portfolios. The goal of the second experiment was to compare and contrast the effectiveness of our framework for financial recommendations with respect to several baselines: specifically, we compared the configurations which emerged as the best ones in the previous experiments (FCV and FCV+Greedy) to the financial domain and the recommendations provided by a human advisor. To further validate our experimental results we also evaluated the performance of the framework against an adaptation of the well-known k-NN algorithm to the financial domain. In the first case, as human recommendations we exploited the the *policy* (i.e. asset-allocation) yield got by each of the 1172 real portfolios stored in the case base from the agreement date to the current one. On the other side, since it was not possible to adopt a user-based CF algorithm due to the sparsity problem (each user signed only one portfolio, and a very little overlap between portfolios existed because investment solutions are always treated as a whole; users do not "buy" single asset classes separately) we proposed an adaption of k-NN algorithm which exploited the distribution of the asset classes in the portfolio to retrieve similar proposals. Specifically, by following this strategy, the mix of the asset classes suggested to the target user was the weighted average of the asset classes suggested to her neighbors. Neighbors were obtained by calculating the overlap between the mix of asset classes in their portfolios

As explained, we calculate the similarity between users according to the overlap between the mix of asset classes in their portfolios, and exploited this value to provide users with recommendations. Results plotted in Figure 10 show that both approaches relying on case-based reasoning significantly outperform both baselines with n=5 and n=10. Specifically, it emerged that the gap gets

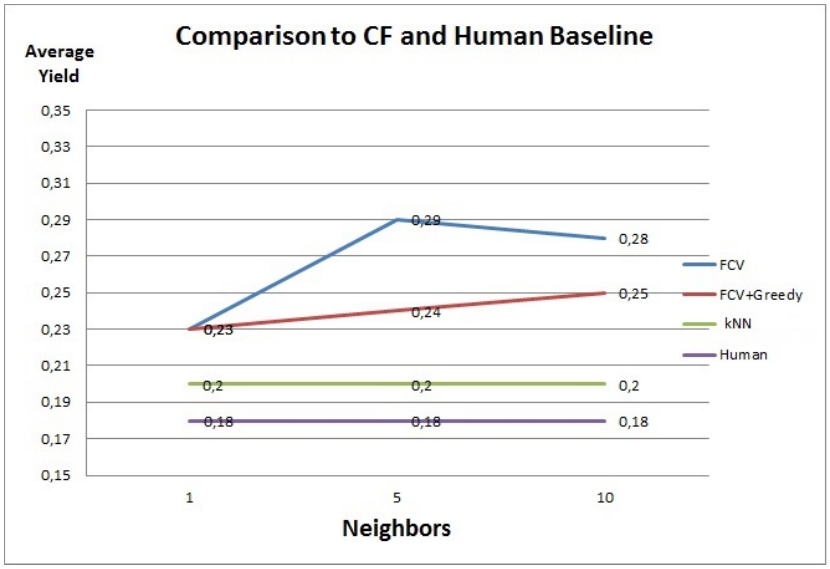


Figure 10: Comparison to k-NN baseline and human advisor recommendations

bigger with *n=5*, even when FCV is compared to FCV+Greedy. This can be justified by the fact that the diversification strategy tends to sacrifice a little of performance (lower average yield) in favour of a greater diversification in terms of investment proposals. Similarly, for n=5 and n=10 FCV+Greedy strategy obtains a statistically significant improvement with respect to both baselines. Finally, it is worth to note that also k-NN is able to generate recommendations better than those provided by human advisors.

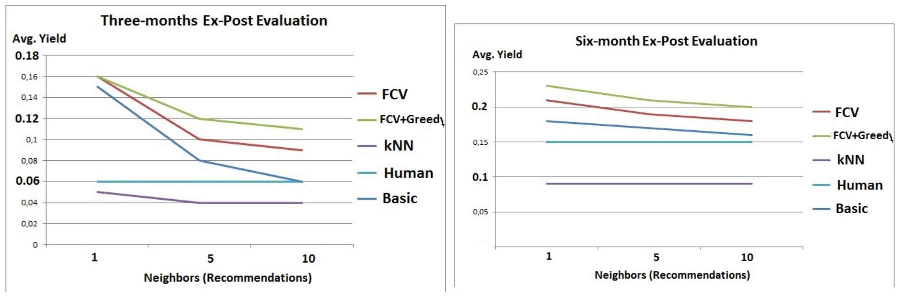
All these outcomes definitely confirm the effectiveness of our recommendation framework, since they show that the proposed approach is able to provide the financial advisor with investment proposals which are better than those he would have suggested to final users.

5.4. Experiment 3: ex-post evaluation

In the *ex-post* evaluation of our framework we compared the *policy* yield gained by the portfolios in two different time intervals: from January 2014 to April 2014 (three-months window) and from January 2014 to July 2014 (six-months window).

In this experiment we simulated that the recommended portfolios were *actually* agreed on January 2014, and we analyzed the *real ex-post perfomance* of the recommendations generated by the framework. To this end, we first calculated the FCV scores by using the historical yield of the asset classes up to January 2014. Next, we generated the recommendations by adopting *FCV* and *FCV+Greedy* strategies. We chose these approaches since they emerged as the best configurations in all the previous experiments. Finally, we calculated the yield obtained by each portfolio from January to April 2014 and from January to July 2014. It is worth to recall that the outcomes of this evaluation are really valuable, since they are based on *real performance* of the portofolios in a *real* time lapse. As shown in Figure 11a and 11b, results produced by this experiment are very interesting: first, differently from Experiment 3, k-NN provides users with a worse average yield than human recommendations. This experimental result further confirms the complexity which characterizes the financial domain and underlines that even such a widespread approach could not provide any benefit to the financial advisor.

On the other side, all the approaches relying on case-based reasoning outperfom the baseline, even with a small gap. For *n=1* and *n=5* the average yield obtained by basic ranking, FCV and FCV + Greedy significantly overcome that obtained by the real portfolios suggested to the users. Anyway, the most interesting outcome is that the best perfoming configuration is not the simple FCV, since the combination of the diversification technique with FCV can further improve the performance of the proposed portfolios. This result suggests that the integration of the approaches can make the framework even more effective. This is due to the fact that a combined strategy can merge the advantages of a ranking based on past performance, as FCV, with an algorithm that may lead to more diverse recommendations. This makes the investment strategy better since the human advisor does not base his investment proposal on a set of very similar portfolios, but rather on a set of *diversified solutions* which is a more stable and effective portfolio, especially when market fluctuations have to be tackled.



(a) January-April 2014 (b) January-July 2014

Figure 11: Ex-post evaluation of the framework

6. Conclusions and Future Work

This paper presented a framework for financial product recommendation relying on case-based reasoning. Our approach integrates a novel strategy based on a Greedy algorithm aiming at diversifying investment proposals. Experiments performed on a dataset of 1172 real users provided several interesting outcomes, since it emerged that the proposed approach can significantly outperform both a baseline represented by a k-NN baseline recommender as well as the recommendations provided by a human advisor. Moreover, an *ex-post evaluation* at three and six months further confirmed these results, since our strategy leads to both diversified and fruitful investment proposals.

As future work, we plan to evolve our recommendation approach in a *conversational* fashion, in order to improve the REVIEW step of the CBR cycle making the advisor able to concretely *discuss* with the recommender system about constrainsts to be relaxed or more specific needs to be matched. Moreover, we will also try to improve the effectiveness of our CBR strategy by incorporating further domain knowledge in the step of retrieving similar cases. An interesting research line may regard the adoption of more expressive representation languages and more complex similarity measures, such as those based on extensions of First-Order Predicate Logic (FOPL) [49]. As regards REVISE strategies, we may adopt generalized linear models (GLM) to rank the candidate proposals on the basis of the prediction of the yield generated by a particular m-dimensional combination of asset classes, which can be learnt on the basis of labeled examples. Finally, we will extend both the representation of the cases by introducing novel features and our ex-post evaluation in order to assess the reliability of our approach in a longer time interval.

References

[1] Vv.Aa., *The Connected Advisor: the Rise of Digital and Social Advice in Wealth Management*, PWC, 2013.

[2] Vv.Aa., *Taking Wealth Management Digital*, Booz and Company Inc., 2013.

[3] B. McDonough, *The Digital Revolution of Wealth Management*, FSC Deloitte Future Leaders Program, 2013.

[4] Vv.Aa., *Closing the Gap: How Tech-Savvy Advisors Can Regain Investor Trust*, Accenture, 2013.

[5] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender systems: an introduction*, Cambridge University Press, 2010.

[6] M.J. Pazzani and D. Billsus *Content-based recommendation systems*, The Adaptive Web, pag. 325-341, Springer, 2007.

[7] Y. Koren and R. Bell, *Advances in Collaborative Filtering*, Recommender Systems Handbook, pag. 145-186, Springer, 978-0-387-85819-7, 2011.

[8] P. Lops, M. de Gemmis and G. Semeraro, *Content-based Recommender Systems: State of the Art and Trends*, Recommender Systems Handbook, pag. 73-105, Springer, 978-0-387-85819-7, 2011.

[9] R. Burke, *Knowledge-based recommender systems*, Encyclopedia of library and information systems, Supplement 32 (69), pag. 175-186, 2000.

[10] W. Lee, *Risk Based Asset Allocation: A New Answer To An Old Question?*, The Journal of Portfolio Management, 37 (4), pag. 11-28, 2011.

[11] A. Aamodt and E. Plaza, *Case-based reasoning: Foundational issues, methodological variations, and system approaches*, AI Communications, 7 (1), pag. 39-59, IOS Press, 1994.

[12] B. Smyth, *Case-based recommendation*, The Adaptive Web, 342-376, Springer, 2007.

[13] J.L. Irving and L. Mann, *Decision making: A psychological analysis of conflict, choice, and commitment*, Free Press, 1977.

[14] J.L. Kolodneer, *Improving human decision making through case-based decision aiding*, AI Magazine, 12(2), pag. 52-70, 1991.

[15] H. Shefrin, *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*, Oxford University Press, 2000.

[16] C. Yu, *Designing A Consumer-Oriented Intelligent Decision Support System For Personalized Financial Planning Services*, 2004.

[17] J. E. Ingersoll, *Theory of financial decision making*, Rowman and Littlefield, 3, 1987.

[18] W.F. De Bondt and T.H. Richard, *Financial decision-making in markets and firms: A behavioral perspective*, National Bureau of Economic Research, 1994.

[19] E.J. Elton and M.J. Gruber, *The rationality of asset allocation recommendations*, Journal of Financial and Quantitative Analysis, 35(1), pag. 27-41, Cambridge University Press, 2000.

[20] J. Merigo and A. Gil-Lafuente, *New decision-making techniques and their application in the selection of financial products*, Information Sciences, 180(11), pag. 2085-2094, Elsevier, 2010.

[21] C. Chuang, *Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction*, Information Sciences, 236, pag. 174-185, Elsevier, 2013.

[22] J.S. Chen, J.L. Hou, S.M. Wu and Y.W. Chang-Chien, *Constructing investment strategy portfolios by combination genetic algorithms*, Expert Systems with Applications, 36 (2), pag. 3824-3828, Elsevier, 2009.

[23] C.C Lin and Y.T. Liu, *Genetic algorithms for portfolio selection problems with minimum transaction lots*, European Journal of Operational Research, 185 (1), pag. 393-40, Elsevier, 2008.

[24] S. Ghazinoory, A.E. Zadeh and A.S. Kheirkhah, *Application of fuzzy calculations for improving portfolio matrices*, Information Sciences, 180(9), pag. 1582-1590, Elsevier, 2010.

[25] A. Fernandez and S. Gomez, *Portfolio selection using neural networks*, Computers and Operations Research, 34(4), pag. 1177-1191, Elsevier, 2007.

[26] P-C. Ko and P-C Lin, *Resource allocation neural network in portfolio selection*, Expert Systems with Applications, 35(1), pag. 330-337, Elsevier, 2008.

[27] N.F. Matsatsinis and E.A. Manarolis, *New Hybrid Recommender Approaches: An Application to Equity Funds Selection*, Algorithmic Decision Theory, pag. 156-167, Springer, 2009.

[28] M-C Chen, C-L Lin and A-P Chen, *Constructing a dynamic stock portfolio decision-making assistance model: using the Taiwan 50 Index constituents as an example*, Soft Computing, 12(11), pag. 1149-1156, Springer, 2007.

[29] C-S Wang and H-L Yang, *A recommender mechanism based on case-based reasoning*, Expert Systems with Applications, 39(4), pag. 4335-4343, Elsevier, 2012.

[30] J. Recio-Garcia, J.A. Gonzalez-Calero and P. Diaz-Agudo, *jcolibri2: A framework for building Case-based reasoning systems*, Science of Computer Programming, 79, pag. 126-145, Elsevier, 2014.

[31] M. Taghavi, K. Bakhtiyari and E. Scavino, *Agent-based computational investing recommender system*, Proceedings of the 7th ACM conference on Recommender systems, pag. 455-458, ACM, 2013.

[32] A. Felfernig and A. Kiener, *Knowledge-based interactive selling of financial services with FSAdvisor*, Proceedings of the National Conference on Artificial Intelligence, 20(3), 1475-1482, 2005.

[33] A. Felfernig, G. Friedrich, D. Jannach and M. Zanker, *An integrated environment for the development of knowledge-based recommender applications*, International Journal of Electronic Commerce, 11(2), 11-34, ME Sharpe, 2006.

[34] I. Gonzalez-Carrasco, R. Colomo-Palacios, J. Lopez-Cuadrado, A. Garcia-Crespo and B. Ruiz-Mezcua, *PB-ADVISOR: A private banking multi-investment portfolio advisor*, Information Sciences, 206, pag. 63-82, Elsevier, 2012.

[35] P. Paranjape-Voditel and U. Deshpande, *A stock market portfolio recommender system based on association rule mining*, Applied Soft Computing, 13(2), pag. 1055-1063, Elsevier, 2013.

[36] R. Burke, *Knowledge-based Recommender Systems*, Encyclopedia of Library and Information Systems, 69(32), A. Kent ed., 2000.

[37] F. Ricci, B. Arslan, N. Mirzadeh, A. Venturini, *ITR: a case-based travel advisory system*, Advances in Case-Based Reasoning, pag. 613-627, 2002.

[38] R. Burke, *The Wasabi Personal Shopper: a case-based recommender system*, Proceedings of the 16th National Conference on Artificial Intelligence and 11th Conference on Innovative Applications of Artificial Intelligence (AAAI-IAAI), pag. 844-849, 1999.

[39] R.G. Ibbotson, *The importance of asset allocation*, Financial Analysts Journal, pag. 18-20, JSTOR, 2010.

[40] R. Baeza Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley, 1999.

[41] R. Xu and D. Wunsch, *Survey of clustering algorithms*, IEEE Transactions on Neural Networks, 16(3), pag. 645-678, 2005.

[42] F. Lorenzi and F. Ricci, *Case-based recommender systems: a unifying view*, Intelligent Techniques for Web Personalization, pag. 89-113, Springer, 2005.

[43] B. Smyth and P. McClave, *Similarity vs. Diversity*, Case-Based Reasoning Research and Development, pag. 347-361 Springer, 2001.

[44] M. Montaner, B. Lopez and J.L. de La Rosa, *Improving case representation and case base maintenance in recommender agents*, Advances in Case-Based Reasoning, pag. 234-238 Springer, 2002.

[45] S. Vargas and P. Castells, *Rank and relevance in novelty and diversity metrics for recommender systems*, Proceedings of the fifth ACM conference on Recommender systems, pag. 109-116, ACM, 2011.

[46] C-N. Ziegler, S.M. McNee, J. Konstan and G. Lausen, *Improving recommendation lists through topic diversification*, Proceedings of the 14th international conference on World Wide Web, pag. 22-32, ACM, 2005.

[47] G. Adomavicius and Y.O. Kwon, *Improving aggregate recommendation diversity using ranking-based techniques*, IEEE Transactions on Knowledge and Data Engineering (TKDE), 24(5), pag. 896-911, IEEE, 2012.

[48] G. Brinson, P. Gary, L. Hood, and G.L. Beebower, *Determinants of portfolio performance*, Financial Analysts Journal, 51(1), pag. 133-138, CFA Institute, 1995.

[49] F. Esposito, D. Malerba and G. Semeraro, *Flexible Matching for Noisy Structural Descriptions*, International Joint Conference on Artificial Intelligence (IJCAI), pag. 658-664, 1991.