Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 181 (2021) 487–494

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2020

Wavelet-based cancer drug recommender system

Liliana Brandão[a,*], Fernando Paulo Belfo[a,b], Alexandre Silva[a,c]

[a]*Coimbra Business School - ISCAC, Polytechnic Institute of Coimbra, Quinta Agrícola, Bencanta, 3045 – 231 Coimbra, Portugal*
[b]*Centro Algoritmi, University of Minho, Campus Azurém, 4800-058 Guimarães, Portugal*
[c]*CEISUC, University of Coimbra, Av. Doutor Dias da Silva 165, 3004 – 512 Santo António Olivais - Coimbra, Portugal*

Abstract

Molecular nature of cancer is the foundation of systematic studies of cancer genomes, providing exceptional insights and allowing treatments advancement in clinic. We combine techniques of image processing for feature enhancement and recommender systems for proposing a personalized ranking of cancer drugs. We use a database containing drug sensitivity data for more than 310.000 IC50, describing response of more than 300 anticancer drugs across 987 cancer cell lines. The system is implemented in Python (Google Colaboratory) and succeed to find best fitted drugs for cancer cell lines. After several preprocessing tasks, regarding drug sensitivity data, two experiments are performed. First experiment uses original DNA microarray images and the second one uses wavelet transforms to preprocess images. Our main goal is to assess the impact of using wavelet transformed DNA microarray images (versus original images) on the proposed framework. The experiments show that, by improving the search of cancer cell lines with similar profile to the new cell line, wavelet transformed DNA microarray images produce better results, not only in terms of evaluation metrics (hit-rate and average reciprocal hit-rate), but also regarding execution time.

Keywords: recommender system; wavelet transform; cancer genome; cancer disease; cell line; DNA; Google Colaboratory; Python

* Corresponding author.
E-mail address: lilianacpbrandao@gmail.com

10.1016/j.procs.2021.01.194

1. Introduction

Recommender systems are automated systems that, in a personalized and meaningful way, lead users to relevant items for them in a wide space of options. Most of their practical applications are web-centric, namely for e-commerce where they engage users by presenting personalized recommendations that best suit their preferences. Nevertheless, recommender systems' potential is much wider. They can be defined as "software tools and techniques providing suggestions for items to be of use to a user" [1]. The terms "suggestions" (or recommendations), "items" and "user" can be understood in a broad sense. The underlying logic behind these systems is anchored on machine learning algorithms that are very versatile and can be applied to many fields. Machine learning has been increasingly used in most diverse domains, either at public sector, such as in fiscal area [2], in education area [3], in the medical field [4] or, at private sector, such as in marketing [5], in media and entertainment industry [6], in events industry [7] and in many other areas, contributing to create new knowledge and helping organizations to define strategies that allow them increase their performance.

This work explores the application of recommender systems in a specific field of medicine domain. The problem under analysis is related to cancer disease. Many research laboratories are testing numerous compounds on cancer cell lines, in order to find the most efficient drugs. Cancer biology is complex and, as being closely related with the physiognomy of each patient, it means that there are some drugs that are more efficient than others in each situation [8]. From this perspective, the problem may be solved with a recommender system supported by machine learning, in which the system aims, given a new cancer cell line (i.e., a new patient), to propose a ranking (i.e., a recommendation) of the most efficient drugs (i.e., items).

Here, a user-based collaborative filtering approach is followed. Consequently, users 'profiles are central to the proposed framework. In the specific context, cancer cell lines are profiled through their corresponding gene expression profile, represented by a DNA (Deoxyribonucleic Acid) microarray. As recognized by [9], there are several microarray systems and methods which differ in several details but produce the same result, an image of spots. On the other hand, images are 2-D spatial signals. Hence, we intent to assess if the prior preprocessing of DNA microarray images, using wavelet transforms, can improve the recommender system performance. To the best of our knowledge, this is the first work that attempts to do it. In practice, such preprocessing represents a shift for the data stored in the DNA microarray image from spatial domain (pixels intensities) to wavelet domain (frequencies). We hypothesize that the representation of the users' profile in a wavelet domain uncovers distinct and discriminating features that improve the search of similar users (a step of vital importance in a user-based recommender system). This is due to the fact that, as stated by [10], wavelet transforms "allow the extraction of richer problem-specific information".

The paper is organized as follows: section 2 provides a brief introduction to the genomic background of the problem; wavelet transforms as a feature extraction technique, namely for DNA microarray images, is described in section 3 along with a theoretical overview of recommender systems and their application to cancer drug recommendation; in section 4 a wavelet-based cancer drug recommender system is proposed; experiments and results are described in section 5 and, finally, section 6 draws the conclusions and discusses future work.

2. Genomics background

In genomics, cell lines refer to the cells capable of renewing themselves in an artificial culture (i.e. under certain laboratory conditions) indefinitely, which makes them ideal for testing new drugs. In what concerns cancer, cell lines are extracted from biopsies (tissues removed from a living body) of patients with different types of tumors (from several body parts depending on the cancer location like lungs or breast).

There are several ways to characterize cell lines, namely using: gene expression, whole-exome sequencing, copy number variation and DNA methylation. However, as shown by [11], gene expression provides "the best predictive power". The gene expression of a cell line is a portrait of the genes' activity contained in that cell and the most common way to measure it is through a DNA microarray.

The DNA microarray slide contains up to tens of thousands of microscopic spots. Each individual spot will be used to measure the activity of a specific gene. This happens during the microarray scanning when the fluorescent intensity of all individual gene spots is stored in an image. A spot with high fluorescence intensity represents a hyperactive gene whereas the absence of fluorescence represents a silent one. Therefore, the DNA microarray image provides a "fingerprint" of the cell line.

Regarding cancer drugs, the efficacy of a compound is usually assessed by the corresponding IC_{50} ("I" for inhibition and "C" for concentration). The IC_{50} is the concentration of the compound required to inhibit the cell growth at 50% [12]. Hence, the lower the IC_{50} value is, the more efficient the compound is.

3. Image processing and wavelet transforms

In the process of microarray scanning, an image of the genes 'activity induced from the fluorescence dye is captured by the scanner. However, "due to the weak fluorescence response, complex biochemical reaction, imperfections in glass slide and photoelectric sensor conversion distortion, etc., the signal of fluorescence probe is inevitably degraded, which leads to serious noise interference in the microarray image" [13].

Wavelet transform is a signal processing technique (also applied to images since they are 2-D spatial signals) that, simultaneously, allows to filter noise and extract more informative features that allow better discriminability of the original signal. The purpose of the wavelet transform is to "transform the signal under investigation into another representation which presents the signal information in a more useful form" [14]. A wavelet is a little wavelike function. During the wavelet transform, a convolution of the signal with a wavelet function happens. This convolution is computed at various locations of the signal (wavelet translations) and for various scales (wavelet dilations). The signal regions where the wavelet overlaps the signal result in large transform values (called wavelet coefficients).

Therefore, several clear structures relating to a specific scale in the wave are detected by shifting the wavelet along the signal [14]. For this reason, wavelet transform has been called a 'mathematical microscope'.

There are several types of wavelets, such as those from Haar, Daubechies, coiflet, and symlet, and it is important to choose the one that best suits our signal and the scope of the analysis. Wavelet transforms come in two distributions: continuous and discrete. The major difference relies in the way how they discretize the scale parameter. The continuous wavelet transform uses exponential scales with a base smaller than 2 (e.g. $2^{\wedge 1/5}$) while the discrete wavelet transform uses exponential scales with the base equal to 2 (i.e., the scales are powers of 2). Hence, continue wavelet transform discretizes scale more finely than the discrete wavelet transform.

However, for image processing, the discrete wavelet transform is the type of distribution usually used, allowing a sparse representation of the signal. Here, coefficients whose value is close to zero may be ignored, remaining only those that have captured important features. The wavelet transform can also comprise several levels of decomposition. In the first level, the signal is decomposed in low and high frequencies regions. The convolution of the wavelet with the low frequency regions result in the so-called approximation coefficients. On the other hand, the convolution of the wavelet with the high frequency regions result in the so-called detail coefficients. In the next level, the approximation coefficients (of the previous level) are again divided into low and high frequency regions. This goes on until it is reached the level of detail that it is needed or until there is no more low and high frequency regions.

The wavelet packet transform is similar to the discrete wavelet transform, however, at each decomposition level, it decomposes not only the approximation coefficients but also the detail coefficients, yielding a higher frequency resolution even in higher frequencies.

In what concerns machine learning classification tasks using images as inputs, the use of wavelets as a preprocessing technique is already in use. For example, in 2010, [15] used the Haar wavelet to extract features of ultrasound liver images for a support vector machine classifier. More recently, [16], working in a classification task using deep neural networks, concluded that the wavelet transform of X-Ray scan improves considerably the performance of the classification network.

We can also find research regarding the application of wavelet transforms to DNA microarray images. For instance, [17] proposed a gene selection method using the discrete wavelet transform on microarray data for cancer classification. Also for a cancer classification task, [18] studied the impact of eighteen different wavelet features extracted from such data.

4. Personalized recommender systems of cancer drugs

This work combines the techniques of image processing for feature enhancement and recommender systems for a personalized ranking of cancer drugs.

The global scope of recommender systems comprises the identification of the need and preferences of users, filtering the huge collection of data accordingly and displaying the best fitted option by using some well-defined

mechanism [1]. Thus, they allow information filtering (in opposition to information retrieval like the one performed by search engines), providing personalized information that is relevant to the user. Recommender systems can have different levels of personalization: from universal (non-personalized recommender systems) to tailor-made (personalized recommender systems). Personalized recommender systems (like the one we propose) make use of the user's specific data to produce recommendations, i.e., one-to-one recommendations.

A possible approach to the recommendation problem is, given a set of unknown items to the user, predict the ratings each item will have and then present the N most (predicted) rated items as recommendations for that user. This is the "prediction version" of the problem [19]. However, it is not mandatory to predict ratings in order to make recommendations. One can simply recommend the top-N most likely relevant items to the user. This is the "ranking version" of the problem and "in many cases it is easier and more natural to design methods for solving the ranking version of the problem directly" [19]. In fact, this is the version commonly adopted in real world problems and the one that it is followed here. In what concerns recommender systems techniques, they comprise several options, namely: content-based filtering, collaborative filtering (user or item-based), model-based methods (or matrix factorization methods), deep learning approaches and hybrid approaches.

For the purpose of this work, we choose a user-based collaborative filtering approach. The assumption on which it is based is that similar users (or, in our case, similar cancer cell lines) share similar behaviours (i.e., similar drug' responses). Therefore, the items (or drugs) that were relevant to existing users will most likely be relevant to a new user similar to them.

Finally, in order to evaluate the recommender system's performance, we use top-N hit-rate and average reciprocal hit-rate. Top-N hit-rate is the fraction of relevant items that are in the top-N recommendation list from all relevant items (the recommended items that are actually relevant items are called hits). As stated by [19], "the disadvantage of the hit-rate is that it gives equal importance to a hit, irrespective of its position in the recommended list". Average reciprocal hit-rate is like top-N hit-rate, but, contrarily to it, it takes into account for where in the top-N recommendation list the hits appear. Its scope is to reward recommended items that match top relevant items.

Most of the research on cancer drug recommender systems has focused on the "prediction version" of the problem, i.e., predicting the exact sensitivity values for the potential drugs. To achieve this goal, a common technique applied is matrix factorization. Matrix factorization solves the recommendation problem by finding latent features that determine the relationship between users and items. For example, [20] use it to project both drugs and cell lines into a latent space (named as "pharmacogenomic space") such that "the dot product between a cell line vector and a drug vector provides the cell line specific drug response". However, a few researchers have also focused on the "ranking version" of the problem. That is the case of [12] whose framework gains "are maximized when the most effective k drugs are the top k recommended drugs", even though, without any drug response being predicted.

5. Proposed framework

The framework proposed here has two main stages: user similarity measurement and cancer drug recommendation. Given a new cancer line (profiled by a DNA microarray image representing its gene expression profile), firstly, a search for the top-N most similar users is conducted and, secondly, using the retrieved information, a personalized cancer drug recommendation is presented, ranking the top-N most effective drugs for the new user.

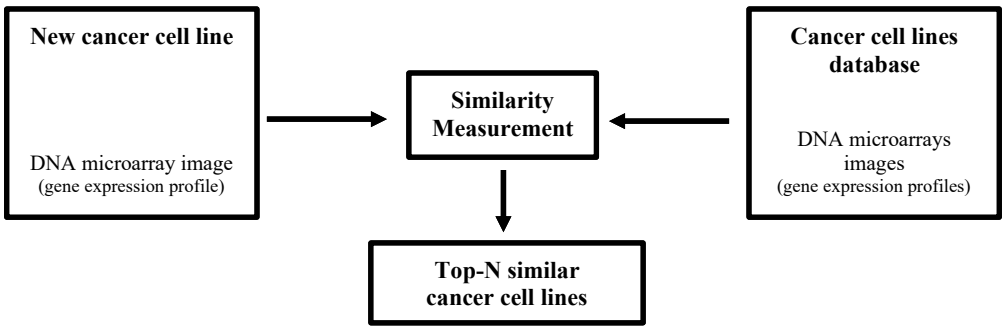

Figure 1. Experiment 1: Experiment without wavelet transforms.

5.1. Stage 1 – Users similarity measurement

The goal of this stage is to find the top-N most similar cancer cell lines regarding the new cancer cell line. This will be done under two experiments: without and with wavelet transforms. Figure 1 presents the framework used at experiment 1, which is done without wavelet transforms. Figure 2 presents the framework at experiment 2, made with wavelet transforms.

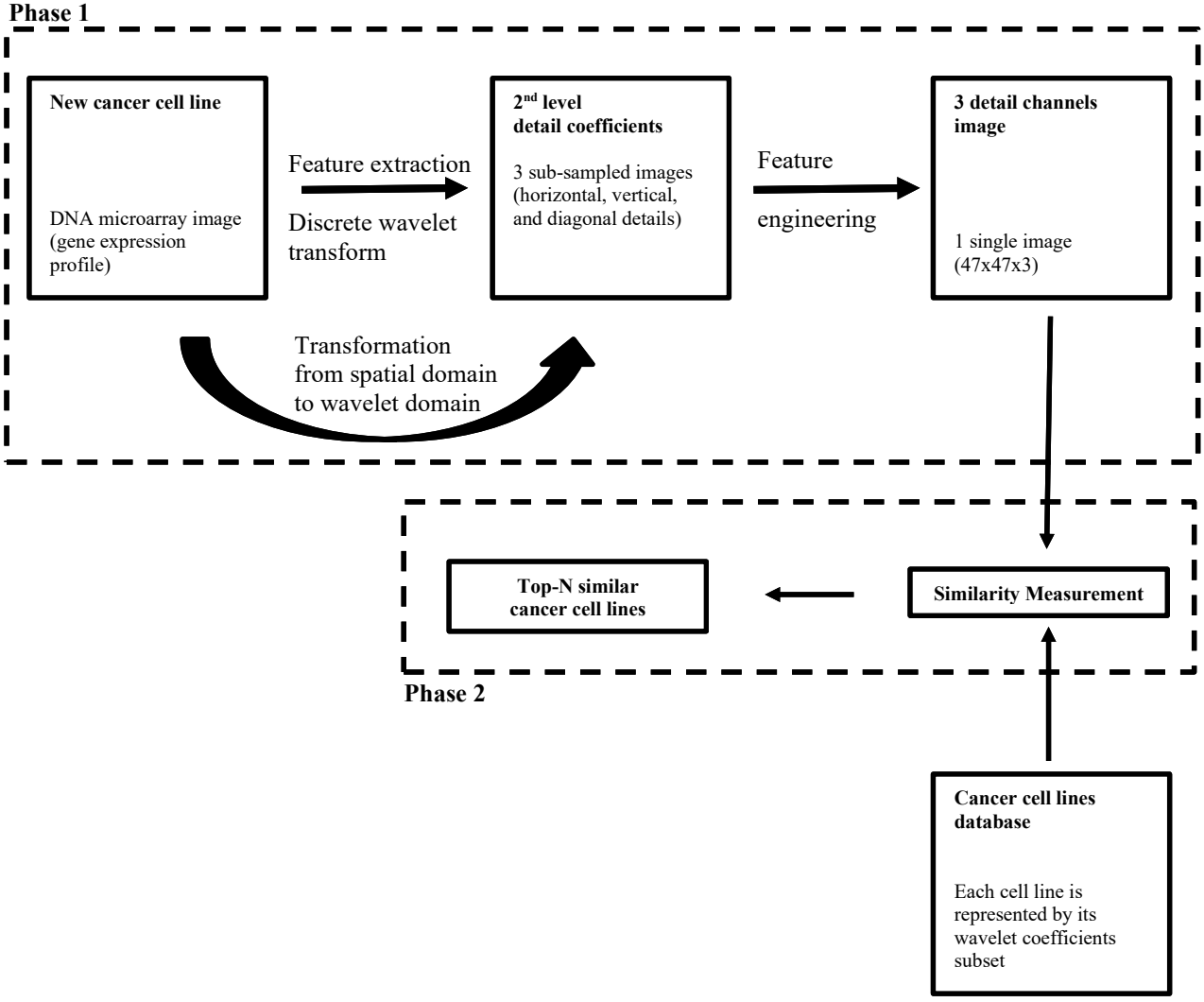

Figure 2. Experiment 2: Experiment with wavelet transforms.

5.2. Stage 2 – Cancer drug recommendation

After finding the most similar users, their drugs' responses are retrieved. Next, the recommendation candidates (i.e., the retrieved drugs) are scored. The score corresponds to the drug's rating (measured by its IC_{50}) weighted by the similarity score between the retrieved cell line and the new cell line. If one drug appears more than once, its scores are added in order to strengthen that fact. The scores are sorted in ascending order. The top-N drugs are then presented as the most likely effective ones for the new cell line (see Figure 3).

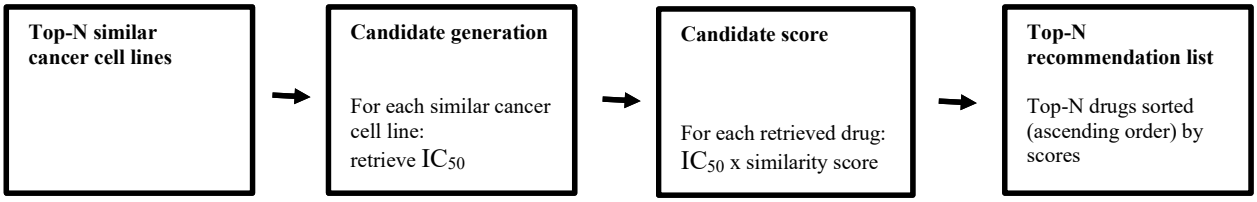

Figure 3. Cancer drug recommendation pipeline.

6. Experiments and results

Experiments are performed on the benchmark dataset called Genomics of Drug Sensitivity in Cancer 1 (release 8.2). The aim is to assess if the use of wavelet transforms on the DNA microarray images contributes positively, or not, to the recommender system's performance. Therefore, the focus lies, not in achieving state-of-the-art results in terms of evaluation metrics (i.e., hit-rate and average reciprocal hit-rate), but on judging the impact of using wavelet transformed DNA microarray images (versus original images) on the proposed framework. To that aim, two experiments are conducted: one using the original DNA microarray images and another one using wavelet transforms to preprocess the images before feeding them to the recommender system.

To measure the similarity between images (i.e., original and wavelet transformed images), Structural Similarity Index [21] is used.

6.1. Dataset

The Genomics of Drug Sensitivity in Cancer (GDSC) database (www.cancerrxgene.org) is one of the largest public resource for information on drug sensitivity in cancer cells. It provides, freely available and without restriction, two datasets - GDSC1 and GDSC2 – which are periodically updated.

The current release 8.2 of the GDSC1 (the one that we will use on the experiments) contains the drug sensitivity data for more than 310 000 IC_{50}, describing the response to more than 300 anticancer drugs across 987 cancer cell lines.

Prior to the experiments, several tasks of preprocessing, regarding the drug sensitivity data, are conducted. First, the IC_{50} concerning cancer cell lines, whose gene expression profiles are not available, are removed. Consequently, only the IC_{50} values of 927 cancer cell lines are kept. Next, to normalize the IC_{50} values into a [0, 1] interval, we follow the method in [22]. Therefore, the closer an IC_{50} value is to zero, the more sensitive the cancer cell line is to the drug whereas the closer the IC_{50} value is to 1, the more resistant the cancer cell line is. The final drug-response matrix (927 cancer cell lines x 345 drugs) has 13.3% of missing IC_{50} values which are set equal to 0.5 (the neutral point in the chosen scale, i.e., the previous interval [0 1]).

For evaluation purposes, the DNA microarray images belonging to the selected 927 cancer cell lines are divided into 4 folders (mutually exclusive but of different sizes): one folder with 852 cancer cell lines and three folders with 25 cancer cell lines each. The three folders of minor size, representing sets of new users in the context of the proposed framework, are evaluated, one at a time. The resulting folder values, for hit-rate and average reciprocal hit-rate, are then averaged and taken as the final result.

6.2. Technological infrastructure

This study uses an environment provided by Google called Colaboratory [23], or just Colab for short. Colab platform requires no setup to use and runs entirely in the cloud, allowing the implementation of machine learning models. This infrastructure allows to write and execute Python code in a browser with zero configuration required and to freely access GPUs (Graphics Processing Unit) and TPUs (Tensor Processing Unit), accelerating the performance of linear algebra computation, which is used heavily in machine learning applications.

6.3. First experiment and its results

Overall, in this experiment, the similarity between images is analysed in a spatial domain (the original domain of the image). In practical terms, this means that the similarity between the new cancer cell line (whose drugs' response are considered to be unknown) and the cancer cell lines of the database (whose drugs' response are known) is measured by applying the Structural Similarity Index (SSIM) between the original DNA microarray image of the new cancer cell line and the DNA microarray images of the cell lines in the database. Afterwards, the drugs' response of the most similar cancer cell line of the database are retrieved. The retrieved IC_{50} values are then weighted by the corresponding SSIM and sorted in ascending order (i.e., from the most to the least effective drug).

The top-20 drugs of the previous ranking are taken as part of the top-20 recommendation list for the new cancer cell line. The final top-20 hit-rate is 11.31 and the average reciprocal top-20 hit-rate is 2.39. Equally remarkable is the execution time of the experiment, approximately 30 hours. These results are set as a baseline for the next experiment.

6.4. Second experiment and its results

Overall, in this experiment, the similarity between images is analyzed in a frequency domain, i.e., between wavelet transformed images.

For most applications, the chosen wavelet type is Haar or Daubechies. Daubechies, although conceptually and computationally more complex than Haar, can pick up details that are missed by Haar. Thus, we choose a Daubechies approach to perform the wavelet transform and assess the hypothesis of an improvement in the recommender system performance through the use of transformed DNA microarray images.

Also, when computing wavelet decomposition, it is possible to use different resolutions (decomposition levels) to convolve the wavelet with the image. Classification using Daubechies 7 with four or five levels of decomposition reported good performance [18]. Therefore, we initially decide to carry out a 4-level decomposition. However, the results suggest that the wavelet transformed image at level 4 is too much compressed resulting in a loss of information. Hence, we proceed instead with a 2-level decomposition.

Although approximation coefficients characterize the major trends contained in the gene expression profiles (i.e., the essential information of the microarray data), we only use the detail coefficients because, as observed by [15], "they have better discriminating capacities and make the classification of two classes of subtle differences possible". In the same line of thought, [24] state that "the purpose of detail coefficients is to detect localized features in one of the gene expression profile".

At the 2nd level of decomposition, and considering only the detail coefficients, three output images are obtained: one in horizontal, other in vertical and another in diagonal directions of the image. These images are then combined, forming a unique 3-detail channels image (similar to a 3-color channels image, using for example, the RGB - Red-Green-Blue color model, but, instead of color channels, the image has the detail channels Horizontal-Vertical-Diagonal). The resulting image is the one used to assess the similarity. This method allows us to preserve spatial patterns (since the coefficients 'positions are kept in the image). The final top-20 hit-rate is 12.21 and the average reciprocal top-20 hit-rate is 2.53. The experiment takes, approximately, 1,5 hours to execute.

7. Conclusions and future work

Recommender systems, comprising the identification of the need and preferences of users, filtering the huge collection of data accordingly and displaying the best fitted options, are becoming more embracing. This study discusses and presents a framework used to implement a recommender system that proposes a personalized ranking of cancer drugs, combining the techniques of image processing for feature enhancement. The proposed framework has a first main stage that consists of measuring the user similarity, and a second main stage, consisting of the cancer drug recommendation. Then, two experiments are conducted. One using the original DNA microarray images and the another using wavelet transforms to pre-process the images before feeding them to the recommender system.

The conducted experiments confirm the initial hypothesis that wavelet transformed DNA microarray images enhance the recommender system performance by improving the search of cancer cell lines with similar profile to the one of the new cancer cell line. In fact, regarding the baseline experiment (that uses the original DNA microarray images), experiment 2 (which uses the corresponding wavelet transformed images) has a top-20 hit-rate and an average reciprocal top-20 hit-rate, 4.5% and 3.89% respectively, higher. Moreover, experiment 2 takes only 5 percent of the

execution time of experiment 1. So, also from a computational point of view, experiment 2 is more efficient and more suitable for a real-world application.

Therefore, we can conclude that properly chosen wavelet transformed DNA microarray images, not only uncover richer information for the users' similarity search (with positive impact, as seen previously, in the recommendation task), but also efficiently compress the DNA microarray images, optimizing computational resources.

Since only 2nd level detail wavelet coefficients were used on this study, future research could investigate the effect of other variants of wavelet transformed DNA microarray images (e.g. simultaneous use of detail and approximation coefficients), with the scope of increasing even further the already existing gap between the evaluation metrics of the two experiments.

References

[1] F. Ricci, L. Rokach, B. Shapira, P. B. Kantor, and F. Ricci, *Recommender Systems Handbook*. Springer, 2011.

[2] A. Seiça, A. Trigo, and F. P. Belfo, "LexiNB - Uma Abordagem Bietápica de Classificação de Sentimentos em Tweets Relacionados com as Autoridades Fiscais Portuguesas," in *Proceedings of the 19.ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI'2019) held in Lisboa, Portugal, 11-12 October 2019. Paper 5.*, 2019.

[3] C. Pimenta, R. Ribeiro, V. Sá, and F. P. Belfo, "Fatores que Influenciam o Sucesso Escolar das Licenciaturas numa Instituição de Ensino Superior Portuguesa," in *18ª Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI 2018) Associação Portuguesa de Sistemas de Informação: Santarém, Portugal*, 2018.

[4] K. J. Cios and G. W. Moore, "Uniqueness of medical data mining," *Artif. Intell. Med.*, vol. 26, no. 1–2, pp. 1–24, 2002.

[5] G. Cui, M. L. Wong, and H.-K. Lui, "Machine learning for direct marketing response models: Bayesian networks with evolutionary programming," *Manage. Sci.*, vol. 52, no. 4, pp. 597–612, 2006.

[6] S. Sereday and J. Cui, "Using machine learning to predict future tv ratings," *Data Sci. Nielsen*, vol. 1, no. 3, pp. 3–12, 2017.

[7] F. Loureiro, A., Lourenço, J., Costa, E., Belfo, "Indução de Árvores de Decisão na Descoberta de Conhecimento: Caso de Empresa de Organização de Eventos," in *VI Congresso Internacional de Casos Docentes em Marketing Público e Não Lucrativo*, 2014.

[8] F. Iorio *et al.*, "A landscape of pharmacogenomic interactions in cancer," *Cell*, vol. 166, no. 3, pp. 740–754, 2016.

[9] J. Serra, "Automatic analysis of DNA microarray images using mathematical morphology," vol. 19, no. 5, pp. 553–562, 2003.

[10] S. Pittner and S. V Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 1, pp. 83–88, 1999.

[11] J. C. Costello *et al.*, "A community effort to assess and improve drug sensitivity prediction algorithms," no. June, 2014.

[12] X. He, L. Folkman, and K. Borgwardt, "Kernelized rank learning for personalized drug recommendation," vol. 34, no. March, pp. 2808–2816, 2018.

[13] Z. Gan *et al.*, "Wavelet Denoising Algorithm Based on NDOA Compressed Sensing for Fluorescence Image of Microarray," *IEEE Access*, vol. 7, pp. 13338–13346, 2019.

[14] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, 2nd ed. Bristol: Institute of Physics Publishing, 2017.

[15] J. Wan and S. Zhou, "Features extraction based on wavelet packet transform for B-mode ultrasound liver images," *Proc. - 2010 3rd Int. Congr. Image Signal Process. CISP 2010*, vol. 2, pp. 949–955, 2010.

[16] A. Rasheed, M. S. Younis, and M. Bilal, "Classification of Chest Diseases using Wavelet Transforms and Transfer Learning," *Micad*, pp. 1–8, 2020.

[17] S. Li, C. Liao, and J. T. Kwok, "Wavelet-Based Feature Extraction for Microarray Data Classification Wavelet-Based Feature Extraction for Microarray Data Classification," no. May 2014, 2006.

[18] L. Nanni and A. Lumini, "Wavelet selection for disease classification by DNA microarray data," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 990–995, 2011.

[19] C. C. Aggarwal, *Recommender Systems - The Textbook*. Springer International Publishing, 2016.

[20] C. Suphavilai, D. Bertrand, and N. Nagarajan, "Predicting Cancer Drug Response using a Recommender System," vol. 34, no. June, pp. 3907–3914, 2018.

[21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[22] C. Properties, M. P. Menden, F. Iorio, M. Garnett, U. Mcdermott, and C. H. Benes, "Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties," no. April, 2013.

[23] Google, "Google Colaboratory," 2020. [Online]. Available: https://colab.research.google.com/. [Accessed: 28-Jan-2020].

[24] Y. Liu and L. Bai, "Find Significant Gene Information Based on Changing Points of Microarray Data," vol. 56, no. 4, pp. 1108–1116, 2009.