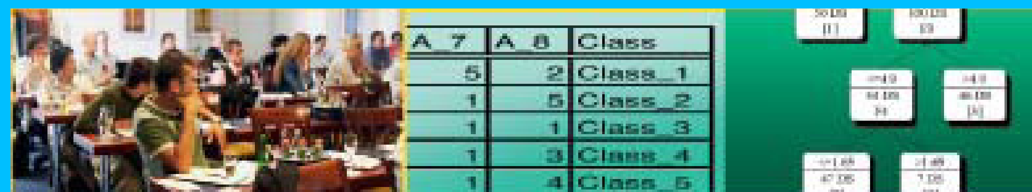
Petra Perner (Ed.)

# Advances in Data Mining



22nd Industrial Conference on Data Mining, ICDM 2022
New York, USA July 13-17, 2022

## Poster Proceedings

ibai - publishing

Advances in Data Mining, Poster Proceedings, ICDM 2022

Petra Perner

ibai-publishing
Prof. Dr. Petra Perner
PF 30 11 38
04251 Leipzig, Germany
E-mail: info@ibai-publishing.org

P-ISSN 1864-9734
E-ISSN 2699-5220
ISBN 978-3-942952-92-7

www.ibai-publishing.org

ISBN 978-3-942952-92-7

9 783942 952927

Petra Perner (Ed.)

# Advances in Data Mining

Applications and Theoretical Aspects

22nd Industrial Conference, ICDM 2022
New York, USA, July 13 – 17 2022
Poster Proceedings

ibai Publishing
www.ibai-publishing.org

Volume Editor

Prof. Dr. Petra Perner
Institute of Computer Vision and Applied Computer
Sciences IBaI

Hertha-Lindner-Str. 10-12
01067 Dresden
E-mail: pperner@ibai-institut.de

The German National Library listed this publication in the
German National Bibliography.
Detailed bibliographical data can be downloaded from http://
dnb.ddb.de.

ibai-publishing
Prof. Dr. Petra Perner
PF 30 11 38
04251 Leipzig, Germany
E-mail: info@ibai-publishing.org
http://www.ibai-publishing.org

P-ISSN 1864-9734
E-ISSN 2699-5220
ISBN 978-3-942952-92-7

# Editorial

The twenty-second event of the Industrial Conference on Data Mining ICDM was held in New York again (www.data-mining-forum.de) running under the umbrella of the World Congress on “The Frontiers in Intelligent Data and Signal Analysis, DSA 2022” (www.worldcongressdsa.com).

At a time when we are still struggling with the corona pandemic, we scientists from different nations have gathered together for a peaceful discourse on an important research focus in the field of data mining and machine learning.

With our conference, we scientists show that we respect the opinions and work of others. That we are ready to consider them peacefully and in friendship under the critical view of the high scientific standards that this conference has.

The conference is an application-oriented conference. In recent years, we have seen a decline in interest in application-oriented research. Probably also because now innovative foundations are focused on oriented research, which requires not so time-consuming work with the experts on site. But we hope that interest will stabilize once we have overcome the consequences of Corona.

The International Program Committee has done an excellent and time-consuming job to select the best papers and provide important guidance on the work of the authors. I would like to thank all the members of the Program Committee for their efforts and that you have contributed with your top-class scientific competence.

The best papers are presented at this conference. The acceptance rate is 33%.

Thank you to all the scientists who have participated in this conference with your excellent work.

A special issue will be done after the conference in the Intern. Journal Transactions on Machine Learning and Data Mining (http://www.ibai-publishing.org/journal/mldm/about.php).

I would also like to thank those scientists who have participated in the conference with their work and have not been successful. Even if we have rejected work, we hope that the indications of the program committee will encourage you to reconsider your work and that you will perhaps face the critical scientific consideration of your work by the international program committee again next year.

The tutorial days rounded up the high quality of the conference. Researchers and practitioners got an excellent insight in the research and technology of the respective fields, the new trends and the open research problems that we like to study further.

A tutorial on Data Mining and a tutorial on Case-Based Reasoning, were held after the conference.

I also thank the members of the Institute of Computer Vision and applied Computer Sciences, Germany (www.ibai-institut.de), who handled the conference as secretariat. We appreciate the help and understanding of the editorial staff at ibai-publishing house, who supported the publication of these proceedings (http://www.ibai-publishing.org/html/proceeding.php).

Last, but not least, we wish to thank all the speakers and participants who contributed to the success of the conference. We hope to see you in 2023 in New York again at the next World Congress on “The Frontiers in Intelligent Data and Signal Analysis, DSA 2023” (www.worldcongressdsa.com), which combines under its roof the following three events: International Conferences Machine Learning and Data Mining, MLDM (www.mldm.de), the Industrial Conference on Data Mining, ICDM (www.data-mining-forum.de, and the International Conference on Mass Data Analysis of Signals and Images in Medicine, Biotechnology, Chemistry, Biometry, Security, Agriculture, Drug Discovery and Food Industry, MDA (www.mda-signals.de), the workshops and tutorials.

July 2022 Petra Perner

# 22nd Industrial Conference on Data Mining ICDM 2022

**www.data-mining-forum.de**

## Chair

Petra Perner
Institute of Computer Vision and Applied Computer Sciences, IBaI, Germany

## Program Committee


| | |
|---|---|
| Ajith Abraham .. | MIR Labs, USA |
| Plamen Angelov | Lancaster University, United Kingdom |
| Antonio Dourado | University of Coimbra, Portugal |
| Stefano Ferilli ... | University of Bari, Italy |
| Warwick Graco. | Analytics Shed, Australia |
| Aleksandra Gruca | Silesian University of Technology, Poland |
| Pedro Isaias....... | The University of New South Wales, Australia |
| Piotr Jedrzejowicz | Gdynia Maritime University, Poland |
| Martti Juhola..... | University of Tampere, Finland |
| Eduardo F. Morales | National Institute of Astrophysics, Optics, and Electronics, Mexico |
| Wieslaw Paja .... | University of Rzeszow, Poland |
| Victor Sheng..... | University of Central Arkansas, USA |
| Iren Todorova Valova | University of Massachusetts Dartmouth, USA |
| Yun Zhao.......... | University of California, USA |

# Table of Content

# Matrix Factorization for Cache Optimization in Content Delivery Networks (CDN)

Adolf Kamuzora[1], Wadie Skaf[1], Ermiyas Birihanu[1], Jiyan Salim Mahmud[1], Péter Kiss[1], Tamás Jursonovics[2], Peter Pogrzeba[2], Imre Lendák[1], and Tomáš Horváth[1]

[1] Data Science and Engineering Department, Faculty of Informatics, Eötvös Loránd University, H-1053 Budapest, Egyetem tér 1-3, Hungary
[2] Deutsche Telekom, Berlin, Germany

**Abstract.** Content delivery networks (CDNs) are key components of high throughput, low latency services on the internet. CDN cache servers have limited storage and bandwidth and implement state-of-the-art cache admission and eviction algorithms to select the most popular and relevant content for the customers served. The aim of this study was to utilize state-of-the-art recommender system techniques for predicting ratings for cache content in CDN. Matrix factorization was used in predicting content popularity which is valuable information in content eviction and content admission algorithms run on CDN edge servers. A custom implemented matrix factorization class and MyMediaLite were utilized. The input CDN logs were received from a European telecommunication service provider. We built a matrix factorization model with that data and utilized grid search to tune its hyper-parameters. Experimental results indicate that there is promise about the proposed approaches and we showed that a low root mean square error value can be achieved on the real-life CDN log data.

**Keywords:** Content Delivery Network · matrix factorization · cache admission · cache eviction · recommender systems

## 1 Introduction

Streaming video services, like Netflix, Disney+, Amazon Prime, and others generate a significant part of Internet traffic today [4]. Content delivery networks (CDNs) are the key components making the delivery of such services with high throughput, low latency a reality. CDNs consist of origin servers and cache servers positioned closest to the customers, i.e., on the edge of the system. The cache servers have limited storage and bandwidth and implement state-of-the-art cache admission and cache eviction algorithms to select the most popular and relevant content for the customers served. In recommender systems, the goal is to suggest similar and popular items to users, given that similar users have requested or rated similar items. Thus, an interest is to predict or determine how the users would rate items that they have not seen. Algorithms such as matrix factorization (MF) could aid this endeavor. The intuition behind matrix factorization is that, with a discovered set of latent features for both users and items, it is possible to predict ratings and thus determine popular and relevant content for the customers.

Our research aims to utilize state-of-the-art recommender system techniques for predicting ratings for cache content in CDNs and their cost reduction potential in real-life content delivery scenarios. Our effort is driven by the needs of a leading European telecommunication service provider, who developed and built its managed CDN to ensure IPTV service delivery. We conducted experiments on an extensive dataset collected in a real-life CDN system. We show that matrix factorization models are viable in the CDN cache environment context.

Our paper is structured into five sections; introduction, related works, methodology section where, the dataset, processing, matrix factorization, evaluation and hyper-parameter tuning are presented. Then results of experiments are presented and finally, the conclusion.

## 2 Related Works

It is already well understood, how CDNs provide the required bandwidth, and thereby ensure reliable, high quality video services [13]. The tremendous network growth faced by CDN operators forces them to continuously innovate and decrease the unit cost of GB delivery. This effort is supported by new hardware generations, state-of-the-art protocols [8], new video encoding standards [3], but mainly affected by the number of request cached on a CDN server: the cache hit rate (CHR). A small improvement in the CHR decreases the resource consumption, reduces the load in deeper caching layers and origin servers and ultimately saves costs.

In spite of the increasing footprint of EDGE cloud services, CDN is the most efficient tool for streaming multimedia contents [13][12]. Adaptive Bitrate (ABR) streaming is a solution which aims to address the challenge for video streaming under unstable network conditions. For ABR streaming, each video is encoded into multiple bitrates, and the customer device selects a bitrate according to the network conditions. This ultimately aims at maximizing users' quality of experience (QoE) [6]. A CDN deploys many servers so that at least some lie in each user's network proximity. CDNs improve service quality and availability by replicating the digital content in those servers. Thus requests to objects (images, video, web content) are handled by nearby servers in case of a cache *hit* (content present in cache). Otherwise, if the requested content is not available in the nearest CDN server, it has to be fetched from the origin server, causing slight delay on the customer side and increasing the load on the origin server. Different researchers claim that the advent of fog computing and machine learning would bring mitigation to some if not all the challenges faced by CDN operators, e.g., origin performance limitations and caching of content [13].

Caching performance improvement in terms of maximizing hit rates is one of the challenges faced in CDN. The authors of reference [1] analyzed the performance of a deployed CDN caching system by deriving its performance metrics using real video on demand (VoD) traces derived from an internet service provider (ISP) CDN caching system. In their analysis, caching system logs that contain a full report of anonymized requests from each user were used, along with the result of the caching operation and time needed to fulfill a particular request. The results of the analyses showed that, first, QoE is optimal if the time to serve is bounded, not if it is the lowest possible one. Also, there is not a straight proportional law between time to serve and resource size. Second, in the case of cache effectiveness, if the cache is serving an optimal number of users: if the per-flow, per-user hit ratio, both normalized to the resource size, are almost identical, then the cache is well performing, and adding more users will not lead to further benefits. Third, it remains the question of cache optimisation by cache admission and cache admission algorithms. Also, the traffic pattern showed to follow a long range dependent distribution (LRD) and not Poisson like, limited or short range dependent [1].

Studies on algorithms for collaborative filtering have proven to be effective in recommendation systems on predicting popular items [9]. While user-based and item-based collaborative filtering methods are intuitive, matrix factorization techniques are effective. This is because, they aid to discover latent factors underlying interactions between users and items [7]. This aspect facilitates especially in the case of determining what content is to be in the cache, whether during cache admission or cache eviction as pointed out by authors in [1].

## 3 Methodology

### 3.1 The Dataset

The analysis was carried out on real-life CDN cache logs collected on multiple hosts in a large European telecommunication service provider. The logs contain detailed, but anonymized requests issued by CDN users, along with the result of the caching operations (hit/miss) and time needed to serve a request. Table 1 provides a brief description of the attributes for the logs.

### 3.2 Data pre-processing

**Explicit vs. implicit feedback** Rating about a product or service on a given scale (e.g., 1 to 5) is often a popular kind of explicit feedback. However, users are usually not easily convinced to give explicit feedback. Other kind of feedback (e.g., views, number of clicks, purchases) is often recorded by systems [10]. In our case, we can extract explicit feedback from the CDN logs pertaining to number of requests for certain content. The cache logs contain a user identifier and the content requested, e.g., live TV channel or video on demand content identifier. This information was used to derive a user-item interaction feature to facilitate our matrix factorization modeling for ratings prediction. Values for the feature were on a log-scale of total requests.

| Feature | Description |
|---|---|
| statuscode | HTTP response status codes |
| contenttype | Indicates the media type of the resource |
| protocol | Http version |
| contentlength | The size of the resource, in decimal number of bytes |
| timefirstbyte | time from request processing until the first byte |
| timetoserv | time needed to process the request |
| osfamily | Client's device Operating System |
| sid | id uniq to a streaming session |
| cachecontrol | Directives for caching mechanisms in both requests and responses |
| uamajor | User-Agent'version, eg. browser's version |
| uafamily | User-Agent,usually client's app |
| devicefamily | Type of the device |
| fragment | video or file fragment identifier |
| path | URL (address of request) |
| timestamp | Arrival time of the request |
| contentpackage | VoD asset identifier |
| coordinates | Long. and lat. of the client based on geoip lookup |
| livechannel | Live TV channel name |
| devicemodel | Client's device model |
| devicebrand | Client's device brand |
| host | Specifies the domain name of the server (for virtual hosting) |
| method | Http request method eg. Get,post |
| manifest | resource which the browser should cache for offline access |
| assetnumber | VoD asset encoding version |
| hit | HTTP request was a cache hit or miss |
| cachename | cache's hostname |
| popname | cache's location |
| uid | id unique to a single user |

**Table 1.** List of log line features

As recommender systems models operate with (user, item) interaction samples, we performed the following mappings from the CDN log data:

$$uid \Longrightarrow user$$

$$livechannel/contentpackage \Longrightarrow item$$

$$accessfrequency \Longrightarrow interaction$$

We grouped user and item Ids and established frequencies for each pair, resulting into a requests feature. We then put the request feature into a log-scale to get the interaction (or preference) feature. We used these features to create a user-item interaction matrix (similar to a utility matrix in recommender systems) as an input to our custom matrix factorization algorithm. Inputs to matrix factorization algorithms of MyMediaLite platform are comma separated value (CSV) files consisting of the userId, itemId and interaction features without column or index names. The results of this approach are depicted in Table 2 showing user, item, frequency and engineered interaction values.

| No | userId | itemId | requests | interaction |
|---|---|---|---|---|
| 0 | 0 | 0 | 16634 | 10 |
| 1 | 1 | 1 | 18038 | 10 |
| 2 | 2 | 2 | 3019 | 8 |
| ... | ... | ... | ... | ... |
| 1978 | 1068 | 2 | 8 | 2 |
| 1979 | 1069 | 13 | 7 | 2 |

**Table 2.** Dataset view for user and item with corresponding interaction

**Splitting Dataset into Train and Test Sets** A random split of the datasets was performed on 70%-30% ratio for training and testing sets respectively (temporal trends were not considered).

### 3.3 Baseline Methods

Matrix factorization is the task of approximating a matrix $X$ by the product of two smaller matrices $W$ and $H$, that is $X \approx WH^T$. In the context of recommender systems, the matrix $X$ is partially observed ratings matrix, $W \in \mathbb{R}^{U \times K}$ where each row $U$ is a vector of users with $K$ latent factors, and $H \in \mathbb{R}^{I \times K}$ describes row $I$ of items, again with $K$ latent factors [11]. Thus, by finding the best $K$ latent features, one should be able to predict a rating with respect to a certain user and item, since the number of features associated with the users match with the number of features associated with the items [7]. Let $w_{uk}$ and $h_{ik}$ be elements of $W$ and $H$, thus the rating of user $u$ to item $i$ is predicted by:

$$\hat{r}_{ui} = \sum_{k=1}^{K} w_{uk} h_{ik} = (WH^T)_{u,i}$$

where $W$ and $H$ are the model parameters which can be learned by optimizing the objective function given a criterion such as root mean square error:

$$\min_{W,H} (r_{ui} - \hat{r}_{ui})^2 + \beta(\|W\|^2 + \|H\|^2)$$

where $\beta$ is a regularization term to mitigate overfitting. For this experiment stochastic gradient descent was used for the optimization process for RMSE.

Initially, we implemented our own matrix factorization class. MyMediaLite recommender system platform's basic matrix factorization algorithm was also utilized. MyMediaLite is free and open-source software which consists of various state-of-the-art recommender system algorithms (e.g., matrix factorization, k-nearest neighbour, most popular item) and an extensive evaluation framework. It runs on the .NET platform, is written in the C# (C-sharp) programming language [10]. MyMediaLite addresses two scenarios in collaborative filtering, rating prediction and item prediction from positive-only feedback. It offers state-of-the-art algorithms, online updates to already trained models, serialization of computed models and various routines for evaluation [5]. Beside basic matrix factorization model described earlier, we also utilized the biased matrix factorization model to factor in user and item biases. The updated model then becomes:

$$\hat{r}_{ui} = \mu + b_u + b_i + \sum_{k=1}^{K} w_{uk} h_{ik}$$

where $\mu$, $b_u$ and $b_i$ are global average, user and item biases respectively. We implemented our experiments in the Python programming language.

### 3.4 Evaluation Techniques

We evaluated our matrix factorization model with RMSE measure. For this metric, a low score is preferred. The RMSE calculation was based on the following equation:

$$RMSE = \sqrt{\frac{\sum_{ui \in D^{test}} (r_{ui} - \hat{r}_{ui})}{|D^{test}|}}$$

The selection of RMSE metric was inspired by the 2009 Netflix competition which was won by a team of researchers, called "Bellkor's Pragmatic Chaos" (10% improvement over CineMatch RMSE) [7].

Evaluation protocols (splitting, candidate selection, metrics) are not easy to get right sometimes. The use of a random seed ensures that the comparison is done on the same "random" split. It can help in aspects such as debugging. We reuse methods/techniques to ensure comparability (more configurations kept fixed thus, less risk of accidental differences)[5]. For this experiment a random seed of 42 was used and GitHub[3] was used as a version control system (scripts and configurations).

### 3.5 Hyper-parameter Tuning

Similarly to other machine learning techniques, the optimal selection of hyper-parameters are very important for matrix factorization. Hyper-parameters for matrix factorization models are latent factors ($K$), the learning rate ($\alpha$), a constant whose value determines the rate of approaching the minimum during parameter search, and regularization ($\beta$), to constrain the model from overfitting.

[3] `https://github.com/CDNResearchProject/Cache_Optimization/blob/main/Notebooks/Rating_Predict_LiveTV.ipynb`

The number of latent features ($K$) is often lower than the number of users and items. The optimal hyper-parameters were found through a grid search. The grid search method attempts all possible combinations of the provided range of hyper-parameter values [2].

## 4 Results

### 4.1 Live-TV Content Analysis

In the initial set of experiments, explicit feedback representation of the data was engineered for the interactions by using request frequencies for CDN content as described in Section 3.2, with Table 2 providing a snapshot of the dataset. The dataset utilized initially pertained to liveTV content.

For this experiment, 5,261,828 cache log lines were used. As described in Section 3.2, the split of train set and test set was on a 70% to 30% ratio respectively. The actual number of records is shown in Table 3.

| LiveTV Datasets | #Attributes | #Records |
|---|---|---|
| Train | 4 | 3,683,280 |
| Test | 4 | 1,578,548 |

**Table 3.** Live TV datasets train-test-split

A custom Matrix Factorization class was written to perform matrix factorization. The main MyMediaLite library ('MyMediaLite.dll') was imported and used in Python.

Table 4 presents the comparison of RMSE of custom implemented matrix factorization and MyMediaLite matrix factorization during preliminary experiments. The results are visualized in the Figure 1 as well.

| No | Method | RMSE |
|---|---|---|
| 1 | Custom matrix factorization class | 0.1981194 |
| 2 | MyMediaLite matrix factorization | 0.1048596 |

**Table 4.** RMSE measures for initial experiments on LiveTV content

### 4.2 Live-TV and VoD Analysis

We performed further experiments with both LiveTV (described in Section 4.1) and video on demand (VoD) content following the same design for initial set of experiments. 3,984,430 cache log lines for VoD were used. After grouping uid with VoD content Ids there were 206,334 rows. This consisted of unique 187,302

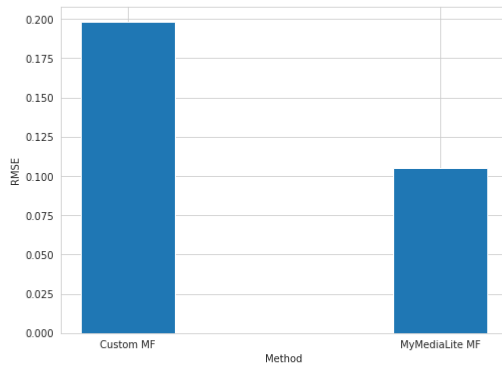



**Fig. 1.** RMSE measures for basic implemented matrix factorization and MyMediaLite matrix factorization on LiveTV content

user ids unique 3940 unique VoD ids. Table 5 shows the actual number of records after the random split

| VoD Datasets | #Attributes | #Records |
|---|---|---|
| Train | 4 | 144,433 |
| Test | 4 | 61,901 |

**Table 5.** VoD datasets train-test-split

Results for biased matrix factorization experiments with LiveTV and VoD datasets are provided in Table 6 and corresponding visualization on Figure 2

| No | Method | RMSE |
|---|---|---|
| 1 | Biased MF (LiveTV) | 0.1628978 |
| 2 | Biased MF (VoD) | 0.1734011 |

**Table 6.** RMSE measures for further experiments on LiveTV and VoD content

We used a custom grid search code compatible with our matrix factorization implementation to perform brute-force search for the best hyper-parameters. MyMediaLite platform contains its own class for grid search which aided in locating the best hyper-parameters. For referencing, the best hyper-parameters that were found in the cross-validation for Experiment 1 are shown in Table 7 and in Experiment 2 in Table 8.

In results of Experiment 1 for Live-TV content analysis, the RMSE measure using algorithm in MyMediaLite is better than our custom implementation of the basic matrix factorization algorithm as shown in Table 5 and Figure 1. In

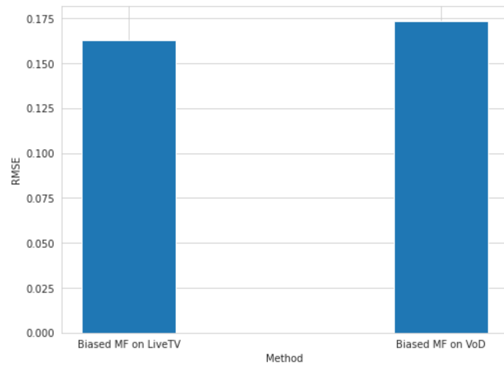



**Fig. 2.** RMSE measures of biased matrix factorization on LiveTV and VoD content

| No | Method | Hyper-parameters |
|---|---|---|
| 1 | Custom class matrix factorization | $K = 46$, $\alpha = 0.3$, $\beta = 0.02$, $iter = 100$ |
| 2 | MyMediaLite matrix factorization | $K = 51$, $\alpha = 0.3$, $\beta = 0.01$, $iter = 100$ |

**Table 7.** Hyper-parameters for LiveTV preliminary experiments

| No | Method | Hyper-parameters |
|---|---|---|
| 1 | Biased MF on LiveTV | $K = 52$, $\alpha = 0.07$, $\beta = 0.05$, $iter = 100$ |
| 2 | Biased MF on VoD | $K = 60$, $\alpha = 0.07$, $\beta = 0.06$, $iter = 100$ |

**Table 8.** Hyper-parameters for LiveTV and VoD content: MyMediaLite platform

Experiment 2 for Live-TV and VoD content analysis, with the biased matrix factorization algorithm of MyMediaLite platform, RMSE measures are comparable in the case of LiveTV and VoD datasets, as shown in Table 6 and visualized in Figure 2. These RMSE scores imply good precision in predicting rating for cache content, thus cost reduction potential in real-life content delivery scenarios. With these matrix factorization models, CDN operators will be equipped with the enhanced capability to predict which live-TV and VoD contents are popular based on ratings (interaction values) and thus to be cached.

## 5 Conclusion

The goal of this research was to show that recommender engines, and matrix factorization in specific, can be utilized in the context of predicting popular content and potentially optimizing cache admission/eviction in a content delivery network setting. We extracted user and content identifiers from CDN logs and generated ratings based on the frequency of requests targeting specific resources (e.g., live TV channels or video on demand content). We used that data to create a matrix factorization model. We showed that a low root mean square error value can be achieved on real-life CDN log data. Future prospects for this work are to leverage CDN logs to further preprocessing, and engineering the ratings in specific, and develop models for item prediction from positive-only implicit feedback which we believe would complement the results presented in this paper.

## References


1. Brilli, L., Fantacci, R., Pecorella, T., Ionta, T.: The effect of video caching on network resource planning—a real-case study. In: In 2016 AEIT International Annual Conference (AEIT), IEEE. pp. 1–6 (2016)
2. Chih-Wei, H., Chang, C.C., Lin, C.J.: A practical guide to support vector classification. In: Gridsearch guide. pp. 1396–1400 (2016)
3. Cianci, P.: High Definition Television: The Creation, Development and Implementation of HDTV Technology. McFarland (2014)
4. CISCO: Cisco annual internet report (2018–2023) white paper (2020), `https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html`
5. Gantner, Z., Rendle, S., Freudenthaler, C., Schmidt-Thieme, L.: MyMediaLite: A free recommender system library. In: Proceedings of the 5th ACM Conference on Recommender Systems (RecSys 2011) (2011)
6. Gao, G., Zhang, H., Hu, H., Wen, Y., Cai, J., Luo, C., Zeng, W.: Optimizing quality of experience for adaptive bitrate streaming via viewer interest inference. IEEE Transactions on Multimedia **20**(12), 3399–3413 (2018)
7. Gábor, T., Pilászy, I., Németh, B., Tikk, D.: Matrix factorization and neighbor based algorithms for the netflix prize problem. In: In Proceedings of the 2008 ACM conference on Recommender systems. pp. 267–274. ACM (2008)
8. Kevin, S., Rahul, U., K, S.R.: Bola: Near-optimal bitrate adaptation for online videos. IEEE/ACM Transactions on Networking **28**(4), 1698–1711 (2020)
9. Mehdi, E., Ricci, F., Rubens, N.: A survey of active learning in collaborative filtering recommender systems. In: Computer Science Review. vol. 20, pp. 29–50 (2016)
10. Steffen, R., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: Bpr: Bayesian personalized ranking from implicit feedback. arXiv preprint **arXiv:1205.2618** (2012)
11. Thai-Nghe, N., Drumond, L., Horváth, T., Nanopoulos, A., Schmidt-Thieme, L.: Matrix and tensor factorization for predicting student performance. In CSEDU **1**, 69–78 (2011)
12. Yu, W., Liang, F., He, X., Hatcher, W., Lu, C., Lin, J., Yang, X.: A survey on the edge computing for the internet of things. IEEE access **6**, 6900–6919 (2017)
13. Zolfaghari, B., Srivastava, G., Roy, S., Nemati, H.R., Afghah, F., Koshiba, T., Razi, A., Bibak, K., Mitra, P., Rai, B.K.: Content delivery networks: state of the art, trends, and future roadmap. ACM Computing Surveys (CSUR) **53**(2), 1–34 (2020)

# Data mining in public health: organizational, ethical, and economic concerns

Marek Navrátil

Masaryk University, Faculty of Economics and Administration, Lipová 41a, 602 00 Brno, Czech Republic

**Abstract.** Data mining (DM) techniques are already showing promise in improving public health across the globe. Before they become more widespread, attention must be paid to the context in which they will operate. This paper scrutinized the relevant literature about the usage of DM and machine learning (ML), concluding that the research has so far failed to encompass the key moral, practical, and fiscal concerns. Using multidisciplinary theoretical concepts, these key recommendations were formulated: Strengthened protection of patients' data must be ensured. The machine-based utilitarian view is best married to the human, deontological approach to form a hybrid model of decision-making. Incomplete public-health data might be a sign of low maturity of current health systems in developed countries.With the aid of public-policy principles, data mining needs to be properly grounded in official national and international institutional structures. Standard tools of assessing medical innovations, such as Health Technology Assessment (HTA), might require some flexibility to achieve a successful economic analysis. Cost-effectiveness calculations are expected to be most convincing in cases with substantial utility, as costs will tend to remain stable. Future research could apply the above principles to specific use cases, public health domains, or countries.

**Keywords:** public health, health economics, data mining, machine learning, ethics, public administration, public policy, health technology assessment, public economics, artificial intelligence.

## 1 Introduction

The complexity of the public health arena lends itself to applications of advanced computational techniques. Efforts have been made to explore the potential of machine learning (ML) and data mining (DM) to achieve better results in various sub-domains of medical care for citizens. The current "golden age" of digital tools used by billions of people around the world makes such applications more promising, now that social media and cell phone data allow for analysis of patient behavioral patterns, epidemiological trends, or effectiveness of various treatments. Before these methods are deployed in a more wide-spread fashion, it is most prudent to explore the various concerns that may arise around their implementation. Unless specified otherwise, the analysis focuses on developed, OECD[1] countries.

The goal of this paper is to outline the context of applications of data mining and machine learning for public health in terms of underlying ethical considerations, practical institutional constraints, and quantifiable economic metrics. The issues are demonstrated based on specific use cases, which have been explored in a variety of settings and from many academic angles. A broader analysis from a public policy/public administration/health economics perspective still lacks in the literature, yet it can serve as a unique glimpse for researchers in applied data mining, decision-makers in health care, and clinical practitioners.

The technological opportunities are many, and so are the associated concerns. This paper shall thus serve not as an exhaustive list, but rather as a starting point to understand the ways in which the former can be successfully leveraged by addressing the latter.

## 2 Methodology

The first goal was to analyze the existing research in this area, using the Scopus and Google Scholar databases. The search was narrowed down to studies focusing on concrete applications of DM and ML techniques in public health to understand the breadth of existing applications. The selected use cases reflected:

- their prevalence in the available literature,
- their interesting features in terms of ethics, organization, and economics.

Another, more narrow set of papers that was collected studied the broader public-facing aspects and implementation principles of these novel approaches. The content of the texts was confronted with theoretical concepts from public policy, health economics, and applied ethics. The combination of theory and practice was meant to serve as a basis for an applied model for technological adoption.

## 3 How to make it right? (Ethical lens)

Once researchers start pondering the ways in which DM could enrich the area of public health, they find themselves confronted with a set of moral questions. These are dependent on deeply ingrained values about the rights of citizens, responsibilities of the state, and optimal allocations of resources. The first ethical question this chapter shall tackle has to do with the actual kinds of data that the technologies will utilize, before considering wider society-wide implications.

[1] Organization for Economic Cooperation and Development.

### 3.1 Individual protections

Data about health is about as sensitive as data gets. On an individual level, precautions need to be in place to prevent breaches that could harm the personal integrity of patients. Such ethical discussions were held during the AIDS epidemic in the 1980s [1], motivated by the risk of revealing the identity of already vulnerable people. Now that more and more data is stored in electronic records, the possibility for an IT hack or internal disclosures seems far more dangerous than in the past. Therefore, investments into cybersecurity and encryption must be made to safeguard this most precious kind of information; the exploration of blockchain-based technologies might be another possible avenue [2]. The control of the individual over his/her health records is paramount also because the power of the state can be used for good as well as for ill – minimizing the potential for violations of human rights in the hands of, e.g., authoritarian actors is, therefore, prudent.

A new promising area of DM applications is digital phenotyping for mental health, which analyzes patterns from the usage of wearable devices and smartphones to estimate various mental states and aid in the prevention of mental health problems [3]. A recent paper by Stanghellini and Leoni [4] explores the ethical terrain, arguing that these technologies cannot give causal accounts for one's health and cautioning about possible breaches. They propose another potential unintended consequence of this innovation: "*Big data may produce a kind of cyber-hypochondria, that is the fear of being or getting sick based on an obsessive monitoring of one's own digitized bodily functions rather than on one's feelings of well-being or ill-being—another example of de-corporealization.*" These factors should be carefully considered when designing these phenotyping techniques.

### 3.2 Limits of collective actions

Looking at the broad implications of rolling out ML and DM technologies on a society-wide level, it is natural to highlight epidemiology and the allocation of resources. The first area has been highlighted by the COVID-19 pandemic, where DM was helpful in predicting incidence in various regions [5]. Even before the latest outbreak, similar techniques were deployed, *inter alia*, in air pollution epidemiology [6]. Understanding the patterns underlying various public health areas lends itself to data-driven budgeting for preventive and reactive health policies alike. The scarcity of available resources in healthcare is a classic ethical conundrum, yet the reactive (and often somewhat reactionary) approaches to medical decision-making still play a key role in many countries. The tide, however, seems to be shifting in the right – preventative – direction, at least in terms of overall trends in the Western world [7]. Given the complex ecosystem of health services with a web of interests, stakeholders, and differential approaches in various fields, the data will be maximally beneficial if used in a sort of "human-machine interface", whereby the algorithms provide the basis on which the decision-makers will choose the best tactics and strategies.

The relationship between the patient-centered and society-centered views of health care can be viewed through the lens of two opposing principles: deontological and utilitarian. The first view establishes a set of principles which must not be violated to adhere to some fundamental societal values. The example in the medical field is the need to provide some basic standard of care to every citizen, regardless of their own sense of responsibility over their health. The second view considers the system as a whole and aims to give preference to those solutions which maximize the overall benefit to society. The medical example here would be the decision to purchase equipment benefitting a small number of people in lieu of a larger group, if the combined benefit-cost-ratio is more advantageous [8]. The current philosophy in the medical sphere seems to favor the deontological approach [9], which is an interesting fact to consider considering possible DM and ML applications; while the data analysis might point to a clear utilitarian calculus, the final decision might have to be altered to be in line with political value-judgments, thus creating a sort of "hybrid" view of public health care. When artificial intelligence (AI) is added to the picture, additional corrections for potential bias will need to be made [10]. In short, the efficiency of machines should, especially in the short term, be married to the wisdom of humans in the medical realm.

## 4 How to make it work? (Organizational lens)

Once the ethical hurdles are cleared, a multitude of practical questions will arise. Public health is inherently, definitionally linked to the public sector. Unlike the private sphere with its profit motives, public administration is prone to inefficiencies and excess bureaucracy. The imperative to serve the public interest as opposed to maximizing revenue also implies several meaningful differences at the intersection between ethics and organization. This chapter shall focus on the practical system readiness as the mining tools wait at the footstep to be formally introduced as a public health tool.

### 4.1 Not quite ready yet

As noted above, ML and DM applications can help alleviate some of the existing problems, but the adoption phase is very likely to be hampered by them. The readiness of individual countries to adopt the new technologies might be evaluated using various maturity models, such as those introduced by Tarhan et al. [11] or Jukić et al. [12]. The latter's contribution can also be used in the possibility to adopt co-creation principles, involving multiple parties from within the public and private sectors to bring about changes in public organizations.

A key issue that will need to be addressed is the availability and interoperability of medical records. Due to the complex health governance encompassing individual physicians, hospitals, government agencies, and international bodies, data is often fragmented and lacks a unified, standardized format [10]. Running algorithms based on incomplete information runs the risk of exacerbating possible errors in decision-making. It follows that effective organization of existing data and clear guidelines for new sources are among the key success factors in the DM adoption process [13]. This is especially true in areas involving truly vast datasets, such as cell phone-enabled diabetes patient reports [14] or genome records [13].

### 4.2 Public policy and legal structures

Formally grounding DM in the legislative frameworks of the adopting countries has the potential of achieving its full potential. This way, democratic legitimacy will be ensured, and implementation rules will be enforceable. Any attempts to bring about the necessary changes in the legal system will be constrained by the specifics of the political system with its actors and their views and interests [16]. The "multiple streams framework" demonstrates the need to align multiple political processes to open a "window of opportunity", allowing for a policy change to take place [17]. In this case, the adoption effort needs to address an existing problem (suboptimal health outcomes and resource allocation), propose a specific solution (building on existing use cases of DM/ML in various public health domains), and engage the key stakeholders (health care professionals as well as politicians on the national level). Technical excellence can only come to fruition when combined with practical political prowess.

Establishing a body tasked with creating standards and conducting oversight can significantly boost the implementation effort. Existing data-focused government agencies can be expanded in this regard to utilize existing expertise. Given the global nature of certain health phenomena (including pandemics), an international unit under the patronage of WHO would ideally provide guidelines for all countries, which in turn need to enforce these on the regional and local levels. The WHO has already taken steps towards this end, giving a hopeful sign [18].

## 5 How to make it cost-effective? (Economic lens)

The question that looms over every decision-maker in health care is that of resources. Data mining is only one of many possible health technologies seeking investment; assessing costs in conjunction with expected utility lies at the core of modern health system management. Coming back to the idea of multiple streams (see section 4.2.), the consideration of DM's value-for-money in its various applications will be key to opening the opportunity window. This chapter introduces a possible tweaking of the currently used methods and lists some principles for future analyses.

### 5.1 Standard tools and non-standard interventions

The practice of Health Technology Assessment (HTA) has become the state-of-the art tool for evaluating the economic aspects of medical procedures, therapeutics, and devices. However, due to the far-reaching implications of the new tools such as AI and DM, the common HTA approaches might fail to capture the full spectrum of costs and benefits. Several papers such as that by Alami et al. [19] provide a possible framework for how the assessment could be performed in the case of AI. Their paper also lays out the necessary investments such as performance/data quality tests, infrastructure upgrades, and human expertise. To make matters more complex, each specific application of DM in different areas of public health will require a separate analysis to deepen the understanding of each case and to be able to prioritize between different implementation projects. As for the input data, trials may need to be run to assess the differential effects of the intervention as opposed to the control groups, aiming to prove that the quantifiable benefits outweigh the costs.

The existing use cases of DM for public health, ranging from breast cancer prevention, mental well-being, viral spread management, to Alzheimer's disease detection, fit very well into the above-mentioned trends of moving into prevention-based interventions. While the investments in the necessary technical infrastructure might be substantial, the long-term impacts are likely to outweigh them (see section 5.2.). For any future HTAs, a societal perspective should be striven for to capture as many effects as possible, although it might not apply to every conceivable case [20]. The instances where a narrower approach (such as payer's perspective) is applied should always be well-argued. In general, a more flexible approach to the design of the assessments might be warranted. The agencies responsible for the management of data tools in health need to provide the necessary support by allocating qualified personnel and providing the necessary data inputs using government sources.

## 5.2 Costs and utility

The benefits stemming from the application of mining methods (in HTA terms known as "utility" [21]) will diverge greatly as they are applied to different medical problems. In some cases, the impact might be marginal (e.g., a time-saving automated data analysis); in others, they might be transformational (e.g., targeted prevention of Alzheimer's disease). The total utility is calculated by the unit impact typically measured by quality-adjusted life years (QALY) multiplied with the number of people affected. Aiming at high-quality or high-quantity interventions can serve as a prioritization principle for future projects. On the other hand, the implementation costs (listed in chapter 5.1) can be expected to be relatively stable across the interventions; they must nonetheless be carefully analyzed within the HTA. It would be beneficial to consider some approaches from cost-benefit analyses, such as "hassle costs" or enforcement costs [22].

Economic analyses must not fail to consider the "softer", second-order impacts of the DM applications for public health. When a new procedure is adopted to track the prevalence of a disease, the analyst's view needs to encompass the potential ramifications described in the previous sections of the paper. Monetizing ethical and practical risks is never easy and always dependent on several assumptions; nonetheless, conducting a careful sensitivity analysis and consulting experts from multiple fields are among the tools available for minimizing errors in the calculations [20].

## 6 Conclusion

The potential of data mining and machine learning techniques to transform the nature of public health is vast. The available literature describes dozens of specific use cases and analyzes past, present, and future challenges. DM and ML implementation was scrutinized by many authors from various countries, yet usually through a narrow lens. This paper connected three aspects key to opening the path for wider adoption, encompassing diverse health systems in terms of underlying values, organization of power, or payment mechanisms. As such, it provides a basis for a model which can be used to guide the adoption of these innovations in the future.

In the area of ethical considerations, a focus was placed on the need to safeguard the data of patients, as well as the need to understand the inherent limitations in the ability to interpret causal patterns from DM analyses. Phenotyping techniques for mental health were used as an example of possible unintended consequences such as "cyber-hypochondria". Two large-scale philosophical approaches to health services provision – deontological and utilitarian – were introduced, concluding that a DM/ML-enabled future might need to marry both approaches, thus creating a hybrid model.

The organizational analysis pointed to the inadequacies in the current system in terms of availability and interoperability of data and listed possible evaluation techniques to assess the level of readiness. A theoretical consideration of public-policy models was introduced to understand the conditions for successful implementation, especially regarding the need to involve the key stakeholders and respond to existing social issues. The value of integrating the DM techniques in national legal frameworks and global structures was emphasized.

Within the economic section, the cost-effectiveness of the DM/ML innovations was considered. The standard HTA approach was considered as a possible strategy for assessing the value-for-money of the technologies, with possible modifications and extra considerations having to be made to capture the large-scale impacts they might have. A basic list of key anticipated costs and utilities of the interventions was compiled – a foundation upon which to build future health economic analyses.

Future research into the area could demonstrate the concepts thus introduced on case studies of a few countries representing different traditions of health care and different political systems. The organizational model could be further elaborated on to provide a concrete checklist or a way of quantifying the maturity of specific countries. Standalone cost-effectiveness studies could follow to assess the economic side of promising projects. The implementation of data-based tools for public health raises multiple concerns, but mainly offers great promise for individual patients as well as whole systems of care.

## References

1. Bayer, F., Fairchild, A. L.: The Genesis of Public Health Ethics. Bioethics 18(6), 473–492 (2014).

2. She, A. et al.. Healthcare Data Breaches: Insights and Implications. Healthcare 8(2), 133 (2020).
3. Martinez-Martin N. et al.: Data mining for health: staking out the ethical territory of digital phenotyping. npj digital Medicine 1, 68 (2018).
4. Stanghellini, G., Leoni, F.: Digital Phenotyping: Ethical Issues, Opportunities, and Threats. Frontiers in Psychiatry 11 (2020).
5. Ahouz, F., Golabpour, A.: Predicting the incidence of COVID-19 using data mining. BMC Public Health 21, 1087 (2021).
6. Bellinger, C. et al.: A systematic review of data mining and machine learning for air pollution epidemiology. BMC Public Health 17(1), 907 (2017).
7. World Health Organization: Lessons from transforming health services delivery: compendium of initiatives in the WHO European Region, https://www.euro.who.int/__data/assets/pdf_file/0014/303026/Compendium-of-initiatives-in-the-WHO-European-Region-rev1.pdf, last accessed 2022/3/16.
8. Mandal, J. et al.: Utilitarian and deontological ethics in medicine. Tropical Parasitology 6(1), 5–7 (2016).
9. McGuire A.. Ethics and resource allocation: An economist's view. Social Science & Medicine 22(11), 1167–1174 (1986).
10. Nicholson Price W. Risks and remedies for artificial intelligence in health care, https://www.brookings.edu/research/risks-and-remedies-for-artificial-intelligence-in-health-care/, last accessed 2022/03/15.
11. Tarhan, A. K. et al.: Maturity assessment and maturity models in health care: A multivocal literature review. Digital Health 6 (2020).
12. Jukić, T. et al.: Organizational maturity for co-creation: Towards a multi-attribute decision support model for public organizations. Government Information Quarterly 39(1), 101623 (2022).
13. Laue H.: Data mining: new opportunities for medicine and public health, https://www.gesundheitsindustrie-bw.de/en/article/dossier/data-mining-new-opportunities-for-medicine-and-public-health, last accessed 2022/03/15.
14. Samways dos Santos B.: Data mining and machine learning techniques applied to public health problems: A bibliometric analysis from 2009 to 2018. Computers & Industrial Engineering 138, 106120 (2019).
15. Yang J. et al.: Brief introduction of medical database and data mining technology in big data era. Journal of Evidence-Based Medicine 13(1), 57–69 (2020).
16. Scharpf, F. W.: Games Real Actors Play: Actor-Centered Institutionalism in Policy Research. 1st edn. Routledge, New York (1997).
17. Kingdon, J: Agendas, Alternatives and Public Policies. 2nd edn. Harper Collins, New York (1995).
18. World Health Organization: Health Data Governance Summit, https://www.who.int/data/events/health-data-governance-summit/introduction, last accessed 2022/03/17.
19. Alami, H. et al.: Artificial Intelligence and Health Technology Assessment: Anticipating a New Level of Complexity. Journal of Medican Internet Research 22(7), e17707 (2020).
20. Culyer, A. J.: "Perspectives" in Health Technology Assessment. AMA Journal of Ethics 23(8), 619–623 (2021).
21. Richardson, J., Schlander, M.: Health technology assessment (HTA) and economic evaluation: efficiency or fairness first. Journal of Market Access & Health Policy 7(1), 1557981 (2019).

22. Randa, A. et al.: Assessing the Costs and Benefits of Regulation. Centre for European Policy Studies, https://www.ceps.eu/ceps-publications/assessing-costs-and-benefits-regulation/, last accessed 2022/03/18.
23. Kolling M. l. et al.: Data Mining in Healthcare: Applying Strategic Intelligence Techniques to Depict 25 Years of Research Development. International Journal of Environmental Research and Public Health 18(6), 3099 (2021).
24. Wu, W. et al.: Data mining in clinical big data: the frequently used databases, steps, and methodological models. Military Medical Research 8, 44 (2021).
25. Karađuzović-Hadžiabdić K. et al.: Evaluation of IBM Watson Natural Language Processing Service to predict influenza-like illness outbreaks from Twitter data. Periodicals of Engineering and Natural Sciences 10(1), 122–137 (2022).
26. Laijawala V. et al.: Mental Health Prediction using Data Mining: A Systematic Review. In. Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST). KJ Somaiya Inst of Engineering & Information Tech, Mumbai (2020).
27. Hemanandhini I. G., Padmavathy, C.: Mental Health Prediction Using Data Mining. In: Smys, S. et al. (eds.) Inventive Computation and Information Technologies. Lecture Notes in Networks and Systems, vol. 336. Springer, Singapore (2022).
28. Pati, A. et al.: IADP: An Integrated Approach for Diabetes Prediction Using Classification Techniques. In: Sahoo, J. P. et al. (eds.) Advances in Distributed Computing and Machine Learning. Lecture Notes in Networks and Systems, vol. 302. Springer, Singapore (2022)

# Exploratory Study on CT 3D Volumetric Image Reconstruction from a Single X-Ray Image Using Deep Learning

Hooi Pin Chew, Bonita VanHeel, Yang Zhu, Alex Fok

University of Minnesota

**Abstract.** Dental radiography (X-ray) is a standard method for detecting dental caries. Because common dental X-ray can only provide limited 2D projected images in the buccal-lingual direction, it is sometimes difficult for dentists to detect accurately the carious regions within a tooth and determine their dimensions in the buccal-lingual plane. Although cone-beam computed tomography (CBCT) is commonly used in the medical field today, and it can create 3D images, it is still not practical for single-tooth diagnosis due to its high cost, high radiation dose and low resolution. Therefore, it is of great interest to be able to generate simulated high-resolution CT images from a single 2D dental X-ray image, which will allow dentists to go beyond 2D X-ray images and visualize in 3D, the extension of caries or other internal defects in patients' teeth. In this paper, we demonstrate a computational model for CT 3D volumetric reconstruction of a molar from a single X-ray image using deep learning methodologies. Specifically, our deep learning model is a Convolutional Neural Network (CNN) based encoder-decoder framework, where the encoder converts higher dimensional information embedded in 2D projections into feature mappings. The decoder then converts the features into 3D entities and the output of the model is in the form of 3D volume data. The training data is a set of coupled data pairs, with each pair consisting of a CT volumetric image and its matching X-ray image. After training, the model can generate a 3D array of intensity values representing a 3D CT volumetric image using a single or multiple 2D X-ray images as input. Preliminary experimental results show that our approach can effectively achieve 3D reconstruction of tooth geometry and internal structure with desirable accuracy from a single X-ray image. However, more training data of higher quality are necessary to improve the accuracy of the predictions.

**Keywords:** Dental Radiographic Image Analysis, Carious Region Detection, Convolutional Neural Network (CNN) based Encoder-Decoder Framework

## 1 Introduction

Dental radiography (X-ray) is a standard method for detecting dental caries. Because common dental X-ray can only provide limited 2D projected images in the buccal-lingual direction, it is sometimes difficult for dentists to detect accurately the carious regions within a tooth and determine their dimensions in the buccal-lingual plane. Although cone-beam computed tomography (CBCT) is commonly used in the medical field today, and it can create 3D images, it is still not practical for single-tooth diagnosis due to its high cost, high radiation dose and low resolution. Therefore, it is of great interest to be able to generate simulated high-resolution CT images from a single 2D dental X-ray image, which will allow dentists to go beyond 2D X-ray images and visualize in 3D, the extension of caries or other internal defects in patients' teeth.

Traditional methods for CT 3D reconstruction from a single 2D projection are statistical shape analysis [1] based. Lamecker et al. [2] and Novosad et al. [3] both used statistical shape analysis techniques to reconstruct CT images using very few X-ray projections. They developed algorithms to assess differences between projections of the X-ray images and the shape of the 3D volume to predict desired 3D output. The limitation for statistical shape analysis-based methods is that they require prior knowledge of the 3D shapes and silhouettes of the object, and it is very sensitive to variations in the input data. Conceivably, 3D reconstruction using deep learning techniques can overcome these limitations if enough data variation is included in the training dataset.

Applying Deep Learning [4] to CT volumetric image reconstruction with ultra-sparse 2D X-ray images is a relatively new field in medical imaging and computer modeling research. Because an X-ray projection is not a purely 2D cross-sectional image, as higher dimensional information is encoded into the image during the projection process, it is feasible to decode anatomical information in the projection direction to achieve volumetric reconstruction from a single X-ray image.

Researchers have demonstrated promising 3D reconstruction results using various deep learning models. Ying et al. [5] presented a solution to reconstruct the 3D volume from 2 X-rays of human chest images. Their unique approach was that for each input data, a separate encoder was used instead of stacking input together as one, which is the mainstream way of dealing with multiple inputs. The 3D reconstruction result was very accurate, but a separate encoder was needed for every input. As an effect, the learning model was huge, and it required large computing resources for training. Henzler et al. [6] introduced a CNN-based encoder-decoder framework [7] that used skip connection and residual learning [8]. They used cranial 2D X-rays and 3D CT of various mammalian species as training data. Only one projection was needed for the network to perform 3D reconstruction with promising results. Song et al. [9] proposed a framework to reconstruct the 3D oral cavity from a single panoramic X-ray (PX) image and prior information of the dental arch. Specifically, they first trained an encoder-decoder based generative model to learn the cross-dimension transformation from 2D to 3D. Then they restored the shape of the oral cavity with a deformation module with the dental arch curve, which could be obtained simply by taking a photo of the patient's mouth. Their proposed technique could restore both the density of bony tissues and the curved mandible surface with the peak signal-to-noise ratio (PSNR) reaching 19 dB.

One common feature among all the work published is that all deep learning models are encoder-decoder network based, where the encoder converts higher dimensional data embedded in 2D projections into feature mappings, the decoder then converts the features into 3D entities, and the output of the model is in the form of CT volumes.

## 2 Methodology

Our prototype was largely based on the approach suggested by Shen et al. [10]. With the highest reconstruction peak signal-to-noise ratio (PSNR > 30 dB) using a single X-ray image as the input, their work demonstrated the best reconstruction result among all previous work we have examined. We implemented the deep learning network architecture introduced in this paper and trained the network using our molar dataset which consisted of X-ray image/volumetric data pairs. The X-ray images were input of the deep learning network and 3D prediction output from the network were compared with corresponding volumetric data for the training validation process.

- Training Data Preparation

The training data set consisted of a set of volumetric data and matching X-ray image pairs.

Our in-house micro-CT scanner was used for collecting volumetric training data. DICOM was the native CT data format, with each scan file having gigabytes in size, which was too large to process for computer systems accessible to us. We thus used the CT-analyzer software to resize the micro-CT scan files to smaller sizes. For our prototyping purpose, the scan files were converted to double precision 128x128x128 raw volumetric images.

The CT data acquisition process was time-consuming; a high-resolution scan could take 45 minutes or more to complete. With limited funding and time, we opted for low-resolution scans so that each scan took about 5 minutes to complete. The low-resolution data was very noisy [Fig.1 a], although applying software filters could remove the noisiness from the data. But this would also introduce uncertainty to material edges [Fig.1 b]. Instead, for this feasibility study, we chose to binarize the data with two levels of grey scale intensities to represent the two dental hard tissues, namely enamel and dentin. The background and hollow parts of a tooth would have intensity level of 0 [Fig.1 c].


**Fig. 1.** Raw CT image (a), filtered image (b) and binarized image (c)

Since we only had volumetric data at hand, we also used the CT-analyzer software to generate simulated X-ray images from original DICOM files. The resulting simulated X-ray images proportionally matched their corresponding volumetric data dimensions.

Currently we have 200 usable volumetric/X-ray image pairs.

- CNN Based Encoder-Decoder Framework

Our prototype deep learning network was implemented based on the model presented by Shen et al. [10], which is publicly available on GitHub. We created our own data manipulation functions to fit our dataset to the model format. The model was a CNN based encoder-decoder framework. Its overall structure is shown in [fig. 2]. Part 1 is the 2D input, which in our case is the simulated X-ray image described in the Training Data Preparation section; Part 2 is the representation network; Part 3 is the transformation network; Part 4 is the generation network; and finally, Part 5 is the 3D volumetric image that represents the predicted 3D shape from the network.

- Representation Network

The representation network is the encoder part of our deep learning network with X-ray images as input. It uses the 2D convolution residual block to assist the deep learning model to learn semantic representations from 2D projections and to convert the original X-ray image input to embedded feature maps. Basic building blocks for this network consist of 2D convolution layers and 2D batch normalization layers. For each step, the 2D convolution layer conducts 2D convolution operations which down-samples the spatial size of the feature map. In addition, to preserve high-dimensional feature representation, the channel number of the feature maps is increased by doubling the number of convolutional filters. A batch normalization layer then follows before feeding feature maps through the rectified linear unit (ReLU) activation layer. Skip connections are implemented to enhance the learning of the feature map at each layer as the skip connection combines the current layer with the feature map learned from the previous layer. The output of the representation network is the feature representation with a size of $4096\times4\times4$ extracted from 2D projections; it will be used as the input source to the transformation network.

- Transformation Network

The transformation network is implemented to connect the representation network with the generation network, which bridges between 2D and 3D feature spaces. This network has three components: 2D feature maps learning, 2D to 3D transformation, and 3D feature cubes learning.

First, as shown in Fig. 2, for feature maps learning, convolution operations with a kernel size of $1\times1$ and ReLU activations are applied across all 2D feature maps in this 2D convolution layer. Next, embedded representations are reshaped from $4096\times4\times4$ to $2048\times2\times4\times4$. In this step, 2D feature maps from previous layer are transformed into 3D feature cubes for subsequent 3D volume generation. Finally, for 3D feature map learning, deconvolution operations with a kernel size of $1\times1\times1$ and sliding stride of $1\times1\times1$ are applied across all 3D feature cubes in this layer to learn 3D representations of the feature map transferred from the previous layer.

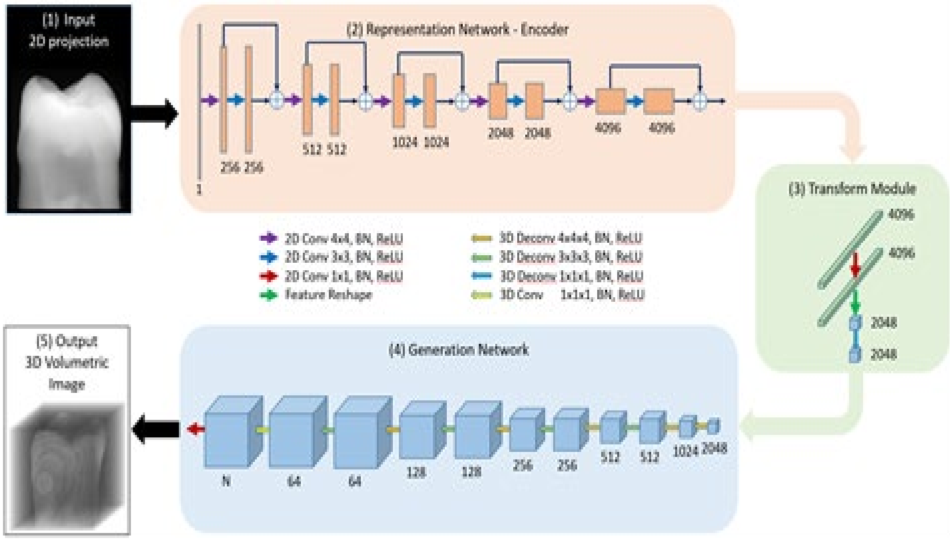



**Fig. 2.** CNN Based Encoder-Decoder Deep Learning Network

- Generation Network

The generation network is the decoder part of our deep learning network. It uses 3D residual deconvolution blocks to convert 3D feature cubes from the transformation network to predicted 3D volumetric images. Basic building blocks for this network consist of 3D deconvolution layers and 3D batch normalization layers. For each step, the 3D deconvolution layer up-samples the spatial size of the feature map. In addition, to transform from a high-dimensional feature domain to a 3D-image domain, the number of feature maps is decreased by halving the number of deconvolutional filters. A 3D batch normalization layer then follows each deconvolution layer before feeding 3D feature maps through the next ReLU activation layer. Finally, the generation network outputs the predicted 3D images fitting the shape of the reconstructed volumetric images.

- Loss Function

Mean squared error (MSE) is used to measure the accuracy of predicted reconstruction results since the 3D reconstruction using deep learning is a regression process. The goal of deep learning networks is to minimize the loss function. In theory, if the prediction is 100% correct, the loss function will have a value of 0. To measure the performance of our deep learning network, we measure the differences between the predicted 3D image and the ground truth using MSE. Here, the ground truth is the volumetric data of the volumetric data/X-ray image pair in the training data set discussed in the Training Data Preparation section, and MSE computes the difference between the prediction and the ground truth for every voxel. If the predicted 3D image matches the ground truth 100%, the value of MSE will be 0.

Since it is impossible to reach the perfect match of 100%, Structural Similarity Index Measure (SSIM) [11] is also used to measure the prediction quality. Unlike MSE, the SSIM models image distortion is a combination of three factors which are loss of orrelation, luminance distortion, and contrast distortion. A 100% match will produce a SSIM value of 1. In our application, a low MSE value combined with a high SSIM value indicates a good prediction result.

## 3 Experiments and Results

We trained the model by feeding a stacked sequence of 2D X-ray projections into the encoder-decoder network iteratively. The output of each iteration was a predicted volumetric image, which was then used to compare with the corresponding ground-truth volumetric image. The loss function between the prediction and the ground truth was the mean squared error (MSE). A training epoch was one complete iteration through the training dataset. At the end of each training epoch, the model was evaluated against the validation dataset which was in the same format as the training dataset.

The model was optimized by iterative stochastic gradient descent, each iteration step was defined as the learning rate, and it was scheduled to decay according to the validation loss in our experiments. An initial learning rate of 2e-5, it would be reduced by a factor 2 if the validation loss remained unchanged for a pre-determined number of epochs. A processing batch containing a group of samples enabled parallel processing using multiple graphics processing units (GPUs); a mini-batch size of 2 were used due to GPU memory limitations. Finally, the best performance network with the smallest validation loss was automatically saved as a training checkpoint and as the final model when an experiment was completed.

Our model could ingest multiple projections per CT volumetric image as inputs. The results presented in this paper are from experiments using a single X-ray projection per CT volumetric image as the training input.

Our network was implemented using the PyTorch [12] library with the Adam optimizer [13] and deployed on two Nvidia RTX5000 GPUs. A complete experiment would run a duration of 10~30 hours depending on the size of a training data set. The typical inference time for a 3D reconstruction from one X-ray image sample was around 0.75 seconds.

We ran experiments with various training database sizes using the same network architecture and training strategy. Table 1 shows results of two of our experiments. Experiment 2 had double the training datasets size of experiment 1, and the model's performance improved slightly because of additional training data available for learning. In Fig. 3, the loss curves of the two experiments indicate that the training process was converging. Fig. 4 shows the 3D reconstruction results from data not included in the training datasets; the same third molar was used for both experiments for the evaluation process. The dimensions of the resulting 3D volumetric image were 128x128x128. Fig. 4 shows cross-sections 20, 40, 50, 60, 70, 80, and 100 of reconstruction results and the ground truth data. Experiment 2 produced slightly better results (mse: 0.0405, psnr: 13.9263, ssim: 0.6601) compared to experiment 1 (mse: 0.0545, psnr: 12.6397, ssim: 0.5781).

**Table 1.** Experiment Parameters and Results

| Exp Number | Training Samples | Validation Samples | Projections per Sample | Iterations | Training Loss | Validation Loss |
|---|---|---|---|---|---|---|
| 1 | 66 | 24 | 1 | 10000 | 0.01703 | 0.05213 |
| 2 | 138 | 47 | 1 | 10000 | 0.0151 | 0.07095 |

## 4 Conclusion

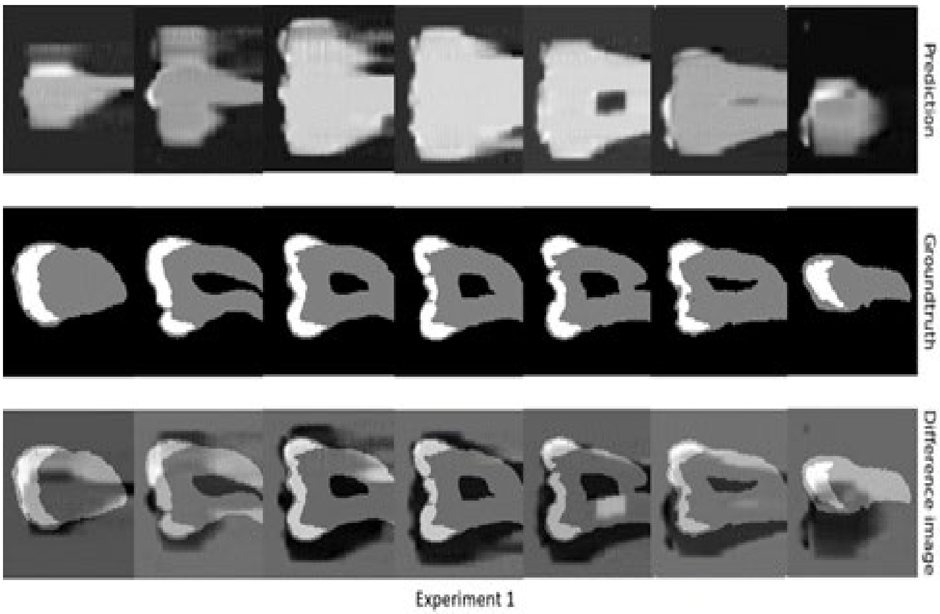
With only 200 samples, we were not able to reconstruct 3D CT images from a single X-ray projection accurately. Nonetheless, this study demonstrated the feasibility of using deep learning methodologies to reconstruct CT volumetric images with ultra-sparse X-rays images as input, even though a large amount of high-quality training data will be needed to achieve improved model performance. As with all deep learning projects, the challenge is the labor-intensive and time-consuming nature of building large high quality training databases.

Also, the success of future study will be constrained by state-of-the-art computing resources. The network architecture discussed in this paper is scalable, but with only 16GB GPU memory available, we are limited to a 128x128x128 reconstruction size.

In addition to building high-quality training datasets, we are also investigating techniques for using principal component analysis methods to reduce data complexities and integrating digitized occlusal surface data into the training process to improve 3D reconstruction quality in the future.

# Appendix


**Fig. 3.** Training and Validation Loss Curves



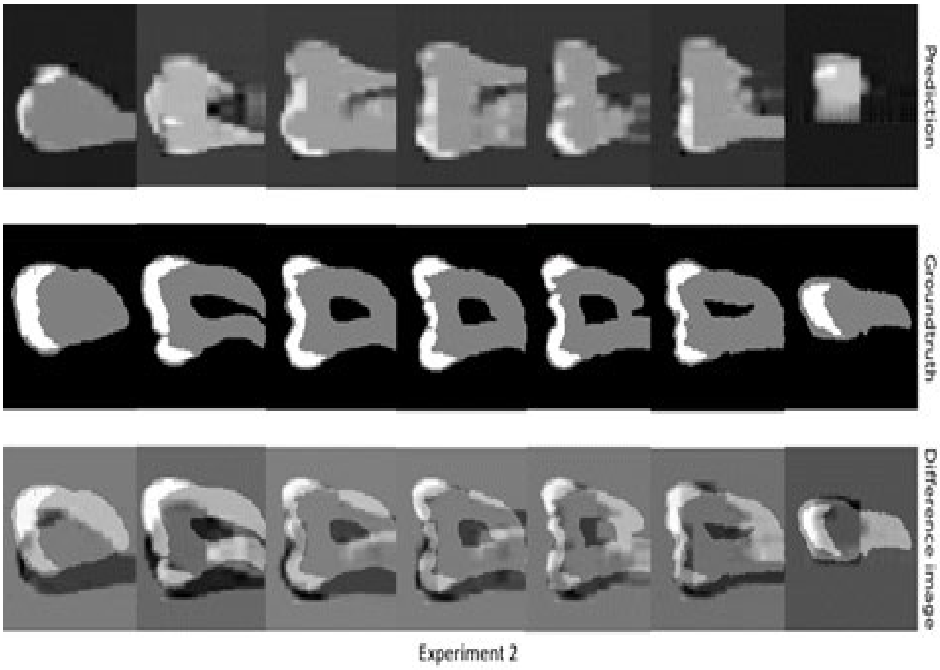



**Fig. 4.** Example 3D Reconstruction Results of a Third Molar

## References


1. Stegmann, Mikkel B., and David Delgado Gomez. “A brief introduction to statistical shape analysis.” Informatics and mathematical modelling, Technical University of Denmark, DTU 15, no. 11 (2002)
2. H. Lamecker, T. H. Wenckebach and H.-C. Hege, "Atlas-based 3D-Shape Reconstruction from X-Ray Images," 18th International Conference on Pattern Recognition (ICPR'06), 2006, pp. 371-374, doi: 10.1109/ICPR.2006.27
3. J. Novosad, F. Cheriet, Y. Petit and H. Labelle, "Three-dimensional (3-D) reconstruction of the spine from a single X-ray image and prior vertebra models," in IEEE Transactions on Biomedical Engineering, vol. 51, no. 9, pp. 1628-1639, Sept. 2004, doi: 10.1109/TBME.2004.827537.
4. Goodfellow, Ian, Bengio, Yoshua, Courville, Aaron. “Deep Learning,” MIT Press, 2016
5. X. Ying, H. Guo, K. Ma, J. Wu, Z. Weng and Y. Zheng, “X2CT-GAN: Reconstructing CT From Biplanar X-Rays with Generative Adversarial Networks,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 10611-10620, doi: 10.1109/CVPR.2019.01087.
6. Henzler, Philipp, V. Rasche, T. Ropinski and T. Ritschel. “Single-image Tomography: 3D Volumes from 2D Cranial X-Rays.” Computer Graphics Forum 37 (2018)

7. Asadi A., Safabakhsh R. (2020) The Encoder-Decoder Framework and Its Applications. In: Pedrycz W., Chen SM. (eds) Deep Learning: Concepts and Architectures. Studies in Computational Intelligence, vol 866. Springer, Cham. https://doi.org/10.1007/978-3-030-31756-0_5
8. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. "Deep Residual Learning for Image Recognition," arXiv:1512.03385 [cs.CV]
9. Weinan Song, Yuan Liang, Jiawei Yang, Kun Wang, Lei He. Oral-3D: "Reconstructing the 3D Bone Structure of Oral Cavity from 2D Panoramic X-ray," arXiv:2003.08413 [eess.IV]
10. Liyue Shen, Wei Zhao, and Lei Xing1,2. "Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning," Nat Biomed Eng. 2019 Nov; 3(11): 880–888.
11. Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, April 2004, doi: 10.1109/TIP.2003.819861.
12. Paszke A. Automatic differentiation in pytorch. Proc. 30th Conference on Advances in Neural Information Processing Systems Autodiff Workshop; 2017.
13. Kingma DP & Ba J Adam: A method for stochastic optimization. Preprint at arXiv:1412.6980 (2014).

# Set Operator Theory Applied to Multidimensional Data

Xiaojin Ye[1,2][0000−0002−4185−0867] and Robert M Haralick[3][0000−0002−2021−4327]

[1] The Graduate Center, CUNY, New York, NY,10016,USA
[2] Fordham University, Bronx, NY, 10458, USA
xye2@gradcenter.cuny.edu
[3] The Graduate Center, CUNY, New York, NY, 10016 USA
rharalick@gc.cuny.edu
http://haralick.org

**Abstract.** It is well known that mathematical morphology plays an important role in image analysis as it enables locating and detecting shapes as well as noise filtering. Our previous papers shows how many of the important properties in mathematical morphology hold in a much more general setting of symbolic or non-numeric sets. This includes the operations of dilation, erosion, opening and closing. This paper extends the relation theory to the power relation theory.

**Keywords:** set operator · indexed relation · indexed power relation · power relation operator · union preserving relation operator · intersection preserving power relation operator · expansive power relation operator · contractive power relation operator · increasing power relation operator · dual power relation operator· power dilation operator · power erosion operator · power relation opening operator · power relation closing operator.

## 1 Introduction

Mathematical morphology plays an important role in image analysis as it enables locating and detecting shapes as well as noise filtering. Our papers ([1–3]) show how many of the important properties in mathematical morphology hold in a much more general setting of symbolic or non-numeric sets. Our generalization has to do with set operators. This includes the set operator for dilation, erosion, opening and closing.

The state of the art of Mathematical Morphology theory is Complete Lattice operators. Such a framework was introduced by Serra and Matheron ([4, 5]) in the eighties, with many contributions for the class of increasing operators. Since then, many researchers have contributed extensions of this theory. In the 1990's, Banon and Barrera developed general lattice operator representations, general set mapping representations.([6, 7]).

The set operator work we did is a specialization of the work that has already been published. Our work is easier to understand since our operators are defined on power sets. All of what we do is restricted to finite sets, and because we work with finite sets we do not need the concepts of infimum and supremum. Therefore, our definition and proofs are easier to understand, and the proofs are more direct.

We have developed many new properties and theorems. Many of these properties and theorems can be extend to the multidimensional data. In this paper, we apply our set operators to multidimensional data, extend the relation theory to the power relation theory. We describe the power relation projection operators, and its inverse.

Multidimensional data is observed as sets of tuples. The observed data can be understood as a sampling of an underlying population of tuples with some of the sampled tuples perturbed in some way. The sampling means that the set of tuples observed is missing the tuples of the population that are not observed. That some of the observed tuples are perturbed means that they may not, in fact, be in the underlying population of tuples and that the information they carry is in essence wrong, just like a noisy value is wrong. The process of inferring order to the real world consists of observing sets of tuples, looking at the patterns in the tuples, and inferring organizing principles called rules which the tuples obey. In physics, for example, the rules are the equations which describe the constraint the underlying variables obey. Organizing principles are simple descriptions of the relationships between and among variables which describe the observed data (fit the data) and which allow predictions to be made about relationships that may not have been apparent from the observed data.

We say that we understand the world described by the observed variables when we have a simply described constraint which the observed tuples of the variables obey. The understanding is embodied in the constraint. We can manipulate the constraint by say fixing some variables of the tuples to a set of values and under these conditions what values the other variables of the tuples have. When we are in a world where the observed variables are symbolic, and the symbol values may not even be partially ordered, we cannot use the mathematics of numbers for the constraint. Rather we have to use relational constraints. And because the principle of scientific inference is to use the simplest description that fits the data, we have to have a set of concepts about relational constraints of various degrees of simplicity. And this is what relational morphology is about.

In section 2, we review the basic definitions and theorems of the different set operators. In section 3, we describe the basic definitions and properties of indexed relation. In section 4, we describe the basic definitions and properties of our set operator on power relation. In section 5, we describe the basic definitions and properties of projection operator on power projection operator. All these high-level theoretical concepts have been completed. Due to the pages limit, we skip the detail proofs of the proprieties and theorems.

## 2 Basic Definitions and Theorems of Set Operators

We do not review papers of mathematical morphology or its extensions to lattices. See([4, 5, 8–29]). In section 1, we review some basic definitions of the different set operators and theories we define in our previous papers.

**Definition 1.** – *A **universal set** is a finite set that contains arbitrary non-numeric elements. We designate whatever universal set we are working with by $U$.[26]*
- *The **power set** of $U$ is the collection of all subsets of $U$, including the empty set, is denoted by $\mathcal{P}(U)$.*
- *A **set operator** $\mathcal{F}$ is a function $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$.*

- *The **operator composition** of a set operator $\mathcal{F}$ with another set operator $\mathcal{G}$ where $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ and $\mathcal{G} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ will be denoted by $\mathcal{G} \odot \mathcal{F}$ and it means first apply $\mathcal{F}$ and then apply $\mathcal{G}$. As appropriate, if we are composing one set operator with another acting on a set A, we may write it as $\mathcal{G}(\mathcal{F}(A))$.*
- *A set operator $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is said to be **increasing** if and only if $A \subseteq B$ implies $\mathcal{F}(A) \subseteq \mathcal{F}(B)$. The operator $\mathcal{F}$ is said to be decreasing if and only if $A \subseteq B$ implies $\mathcal{F}(A) \supseteq \mathcal{F}(B)$.[21]*
- *An operator $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is said to be **expansive** if and only if $A \subseteq \mathcal{F}(A)$. The operator $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is said to be **contractive** if and only if $\mathcal{F}(A) \subseteq A$.[21]*
- *Let $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$. $\mathcal{F}$ is called an **union preserving operator** if and only if*

$$\mathcal{F}(A \cup B) = \mathcal{F}(A) \cup \mathcal{F}(B)$$

- *Let $\mathcal{G} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$. $\mathcal{G}$ is called an **intersection preserving operator** if and only if*

$$\mathcal{G}(A \cap B) = \mathcal{G}(A) \cap \mathcal{G}(B)$$

- *An operator $\mathcal{D} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is called a **set dilation operator** on U if and only if $\mathcal{D}$ is expansive: and union preserving.*
- *An operator $\mathcal{E} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is called a **set erosion operator** on U if and only if $\mathcal{E}$ is contractive and intersection preserving.*
- *An operator $\mathcal{T} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is called a **closing operator** on U if and only if $\mathcal{T}$ is*
  - *Expansive: $A \subseteq \mathcal{T}(A)$*
  - *Increasing: $A \subseteq B$ implies $\mathcal{T}(A) \subseteq \mathcal{T}(B)$*
  - *Idempotent: $\mathcal{T}(\mathcal{T}(A)) = \mathcal{T}(A)$*

  *An operator $\mathcal{Q} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is called an **opening operator** on U if and only if $\mathcal{Q}$ is*
  - *Contractive: $\mathcal{Q}(A) \subseteq A$*
  - *Increasing: $A \subseteq B$ implies $\mathcal{Q}(A) \subseteq \mathcal{Q}(B)$*
  - *Idempotent: $\mathcal{Q}(\mathcal{Q}(A)) = \mathcal{Q}(A)$*

  *[21][30][31]*
- *Let $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ be a union preserving operator. Then we define its **inverse** by*

$$\mathcal{F}^{-1}(A) = \bigcup_{\{X \mid \mathcal{F}(X) \subseteq A\}} X$$

- *Let $\mathcal{G} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ be an intersection preserving operator. Then we define its **inverse** by*

$$\mathcal{G}^{-1}(A) = \bigcap_{\{X \mid \mathcal{G}(X) \supseteq A\}} X$$

- *Let $\mathcal{F} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$. An operator $\mathcal{G} : \mathcal{P}(U) \rightarrow \mathcal{P}(U)$ is called the **dual operator** to $\mathcal{F}$ if and only if $\mathcal{G}(A) = \mathcal{F}(A^c)^c$.*

Table 1: Shows a summary of dual operators' propositions

| Operator $\mathcal{F}(A)$ | Dual Operator $\mathcal{G}(A) = \mathcal{F}(A^c)^c$ |
|---|---|
| Expansive | Contractive |
| Union preserving | Intersection preserving |
| Set Dilation | Set Erosion |
| closing | opening |

– *Let $\mathcal{F} : \mathcal{P}(U) \to \mathcal{P}(U)$ and $\mathcal{G} : \mathcal{P}(U) \to \mathcal{P}(U)$. The operator $\mathcal{F}$ is called the* ***left adjoint*** *to the operator $\mathcal{G}$ and $\mathcal{G}$ is called the* ***right adjoint*** *of the operator $\mathcal{F}$ if and only if for every pair of subsets $A, B \subseteq U$,*

$$\mathcal{F}(A) \subseteq B \text{ if and only if } A \subseteq \mathcal{G}(B)$$

In mathematical morphology, dilation and erosion can define the closing and opening. Similarly, for any set, the union preserving operator and the intersection operator can construct the closing and opening. We define 8 theorems to construct the closing and opening operators by using the set operators and the inverses we defined.

**Theorem 1.** *Let $\mathcal{F} : \mathcal{P}(U) \to \mathcal{P}(U)$ be a union preserving operator. Then $\mathcal{F}^{-1} \odot \mathcal{F}$ is a closing operator.*

**Theorem 2.** *Let $\mathcal{G} : \mathcal{P}(U) \to \mathcal{P}(U)$ be an intersection operator. Then $\mathcal{G}^{-1} \odot \mathcal{G}$ is an opening operator.*

**Theorem 3.** *Let $\mathcal{D} : \mathcal{P}(U) \to \mathcal{P}(U)$ be a set dilation operator. Then $\mathcal{D}^{-1} \odot \mathcal{D}$ is a closing operator.*

**Theorem 4.** *Let $\mathcal{E} : \mathcal{P}(U) \to \mathcal{P}(U)$ be a set erosion operator. Then $\mathcal{E}^{-1} \odot \mathcal{E}$ is an opening operator.*

**Theorem 5.** *Let $\mathcal{D} : \mathcal{P}(U) \to \mathcal{P}(U)$. If $\mathcal{D}$ is a set dilation operator, and its dual $\mathcal{E}(A) = \mathcal{D}(A^c)^c$ is a set erosion operator. Then $\mathcal{E}^{-1} \odot \mathcal{E}$ is an opening operator.*

**Theorem 6.** *Let $\mathcal{E} : \mathcal{P}(U) \to \mathcal{P}(U)$. If $\mathcal{E}$ is a set erosion operator, and its dual $\mathcal{D}(A) = \mathcal{E}(A^c)^c$ is a set dilation operator. Then $\mathcal{D}^{-1} \odot \mathcal{D}$ is a closing operator.*

**Theorem 7.** *Let $\mathcal{F} : \mathcal{P}(U) \to \mathcal{P}(U)$ be a set dilation operator. Let $\mathcal{G} : \mathcal{P}(U) \to \mathcal{P}(U)$ be its dual, a set erosion operator. If $\mathcal{F}$ is the left adjoint to $\mathcal{G}$, then $\mathcal{G} \odot \mathcal{F}$ is a closing operator.*

**Theorem 8.** *Let $\mathcal{F} : \mathcal{P}(U) \to \mathcal{P}(U)$ be a set dilation operator. Let $\mathcal{G} : \mathcal{P}(U) \to \mathcal{P}(U)$ be its dual, which is a set erosion operator. If $\mathcal{F}$ is the left adjoint to $\mathcal{G}$, then $\mathcal{F} \odot \mathcal{G}$ is an opening operator.*

We developed many set operators, including increasing operators, decreasing operators, expansive operators, contractive operators, union preserving operators, intersection preserving operators, set dilation operators and set erosion operators, dual operators, and adjoint operators.

The composition of the set erosion operator with its dual, the set dilation operator is contractive, increasing and idempotent, it is called an opening operator. It is an operator that has the property of producing open sets. With such an operator, we can take a set that has paper shred garbage nearby or touching the set and operate on it to remove the garbage. This happens for arbitrary sets just in the analogous way that the opening operator of mathematical morphology produces that subset of the original set where every point in the subset is an interior point.

The composition of the set dilation operator with its dual, the set erosion operator is expansive, increasing and idempotent, is an operator called a closing operator. With such an operator, we can take a set that has even many small holes and operate on it to produce a set for which the holes are eliminated. This happens for arbitrary sets just in the analogous way that the closing operator of real analysis produces a set that includes the original set plus all its limit points.

Additionally, the original set dilation operator, has for its dual the set erosion operator, so it is the case that the closing operator has for its dual the opening operator. The composition of a closing operator with an opening operator is idempotent. Similarly the composition of an opening operator with a closing operator is idempotent. With the above properties, there is enough structure in a set dilation operator and its dual, to produce operator compositions that can fill holes and gaps in the observed data and eliminate paper shred garbage, thereby changing the observed data set into one whose pattern is closer to the pattern in the underlying population from which the observed data set was sampled.

We can apply our set operators to multidimensional data using the relation theory, the power relation theory and power relation projection operators, and its inverse.

## 3 Indexed Relation

We assume a standard form of observation. There are $N$ variables $X_1, \ldots X_N$ and each observation is an $N$-tuple $(x_1, \ldots, x_N)$ lying in the Cartesian product set $L_1 \times L_2 \times \cdots \times L_N = \bigtimes_{n=1}^{N} L_n$. Thus are observation set is an $N$ relation $R \subseteq \bigtimes_{n=1}^{N} L_n$.

What we need is the ability to manipulate relations and operate on relations. Since an $N$ relation is a set, the set operations of union, intersection, and complementation are usable. Now we extend this set of operations. Our first new concept is the $(I, J)$ relation join. In this definition, $\{I, J\}$ is a cover of $\{1, \ldots, N\}$ the index set of the variables. There is no constraint about whether or not $I \cap J = \emptyset$. The idea is that the join constructs tuples by taking part of the tuples in R and joining them to some of the tuples in $S$ where the tuples match on their common indexes. We can think of the construction operation by starting with the components of a tuple in $R$, match the indexed components specified by $I \cap J$ to the corresponding components of a tuple of $S$, and then construct the remainder of the tuple with the values of the components indexed by $J - I \cap J$ of the matched tuple from $S$ .

Let $X_1, \ldots, X_N$ be the $N$ variables associated with a relation. Let $L_i$ be the set of possible values variable $X_i$ can take. We can define a data set or knowledge constraint relation $R$.

**Definition 2.** **(Relation)** *R is called a relation on range sets $L_1, \ldots, L_N$ if and only if*

$$R \subseteq \bigtimes_{i=1}^{N} L_i$$

*Where $\bigtimes_{i=1}^{N} L_i = L_1 \times L_2 \times \cdots \times L_N$ is a Cartesian product.*

We will be working with many relations associated with different and overlapping variable sets and therefore over different domains. For this purpose we will carry an index set along with each relation. The index set indexes the variables associated with the relation. An index set is a totally ordered set.

**Definition 3.** **(Index set)** *$I = \{i_1, \ldots, i_N\}$ is an index set if and only if $i_1 < i_2 < \cdots < i_N$.*

**Definition 4.** *The order of a relation $R \subseteq \bigtimes_{i=1}^{N} L_i$ is N.*

For a natural number $N$, we let $[N]$ designate the set $\{1, \ldots, N\}$, and $|A|$ designate the number of elements in the set $A$. Next we will define the relation with respect to an index set.

**Definition 5.** **(Indexed Relation[32])** *Let $X_1, \ldots, X_N$ be variables with respective range sets $L_1, \ldots, L_N$. For any $I \subset [N]$, an indexed relation with respect to the range sets $\{L_n\}_{n=1}^{N}$ is a pair $(I, R)$ where*

$$R \subseteq \bigtimes_{i \in I} L_i$$

*Here, $\bigtimes_{i \in I} L_i = \bigtimes_{i=1}^{N} L_i$*

The Index Relation is an instance of a set operator, so we defined the index relation as the relation operator.

**Definition 6.** **(Indexed N-ary Relation)** *If I is an index set with $|I| = N$ and $R \subseteq \bigtimes_{i \in I} L_i$, then we say $(I, R)$ is an indexed N-ary relation on the range sets indexed by I. We also say that $(I, R)$ has dimension N.*

**Definition 7.** **(Relation Composition)** *Let $(I, R)$ be an indexed relation with respect to $\{L_i\}_{i \in I}$. Let $h = (h_i)_{i \in I}$ be a tuple of functions $h_i : L_i \to M_i$. The relation composition of $(I, R)$ with h, with respect to $\{M_i\}_{i \in I}$, the indexed relation $(I, S)$, where $S = \{(s_1, \ldots, s_{|I|}) \in X_{i \in I} M_i \mid \text{ for some } (r_1, \ldots, r_{|I|}) \in R, s_i = h_i(r_i), i \in [I]\}$. We denote this composition by writing $(I, R) \circ h = (I, S)$*

## 4 Set Operator on Power Relation

**Definition 8. (Power Relation)** *A set operator $\mathcal{S} : \mathcal{P}(L) \to \mathcal{P}(L)$ maps a subset to a subset. For relation $R \subseteq \bigtimes_{n=1}^{N} L_n$, every element is a tuple from $\bigtimes_{n=1}^{N} L_n$.*
*A power relation $\mathbb{R}$ is defined by:*

$$\mathbb{R} \subseteq \bigtimes_{n=1}^{N} \mathcal{P}(L_n)$$

Since a relation is a subset of a Cartesian product set, we can say that the power relation is a subset of Cartesian product of power sets. That is, the power relation is defined on a set of tuples whose components are subsets.

The first component is a subset of $L_1$, the second component is a subset of $L_2$,..., the $N$th component is a subset of $L_N$. The power set relation operators consist of $N$ set operators, $S_1$,$S_1$, ..., $S_N$. $S_1$ operate on the fist component, $S_2$ operate on the second component, ..., $S_N$ operate on the $N$th component. In the general case, all components have different operators. So, if there are 100 components, there 100 set operators.

Generalizing the relation to the power relation, we can apply our set operator properties in $N$ dimensions. To do the four operations dilation, erosion, opening and closing, we can do it in the case where we are dealing with an $N$ dimensional relation whose tuple has components which are subsets.

**Definition 9. (Power Relation Operator)**
*Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$,*
*A power relation operator $S$ is defined by*

$$S(\mathbb{R}) = \left\{ \begin{pmatrix} W_1 \\ \ldots \\ W_N \end{pmatrix} \in \bigtimes_{n=1}^{N} \mathcal{P}(L_n) \;\middle|\; \textit{for some} \begin{pmatrix} V_1 \\ \ldots \\ V_N \end{pmatrix} \in \mathbb{R}, \begin{pmatrix} W_1 \\ \ldots \\ W_N \end{pmatrix} = \begin{pmatrix} S(V_1) \\ \ldots \\ S(V_N) \end{pmatrix} \right\}$$

Now we can extend the set operator we defined to the power relation operator.

A power relation operator is called union preserving, if it's union preserving on each component.

**Definition 10. (Union Preserving Power Relation Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called union preserving power relation operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each union preserving.*

A power relation operator is called intersection preserving, if it is intersection preserving on each component.

**Definition 11. (Intersection Preserving Power Relation Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called intersection preserving power relation operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each intersection preserving.*

A power relation operator is called expansive, if it is expansive on each component.

**Definition 12. (Expansive Power Relation Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called expansive power relation operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each expansive.*

A power relation operator is called contractive preserving, if it is contractive on each component.

**Definition 13. (Contractive Power Relation Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called contractive power relation operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each contractive.*

A power relation operator is called increasing, if the set operators for each component are increasing.

**Definition 14. (Increasing Preserving Power Relation Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called increasing power relation operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each increasing.*

We restrict ourselves to operators which have duals. Two power relation operators are dual if is it that dual for each component.

**Definition 15. (Dual Power Relation Operators)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ and $Z = (\mathcal{Z}_1, \ldots, \mathcal{Z}_N)$, where $\mathcal{Z}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be power relation operators. Then $S$ is called the dual operator to $Z$ if and only if $\mathcal{S}_n(A) = \mathcal{Z}_n(A^c)^c$ for all $n$.*

A power dilation is a power relation operator in which the operator is union preserving and expansive, that is, it is union preserving on each component and it is expansive on each component.

**Definition 16. (Power Dilation Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called a power dilation operator if and only $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each union preserving and expansive.*

A power erosion is a power relation operator in which the operator is intersection preserving and contractive, that is, it is intersection preserving on each component and it is contractive on each component.

**Definition 17. (Power Erosion Operator)** *Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called power erosion operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each intersection preserving and contractive.*

**Definition 18. (Power Relation closing Operator)**
*Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then $S$ is called a power relation closing operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each contractive, increasing, and idempotent.*

**Definition 19. (Power Relation Opening Operator)**
*Let $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$, where $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ be a power relation operator. Then, $S$ is called a power relation opening operator if and only if $\mathcal{S}_1, \ldots, \mathcal{S}_N$ are each expansive, increasing, and idempotent.*

The opening is a set erosion operator followed by its dual set dilation operator. Extending this to power relation, a power relation opening operator is a power erosion followed by its dual power dilation operator.

**Theorem 9.** *Let* $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$*, where* $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ *be a power dilation operator. Let* $Z = (\mathcal{Z}_1, \ldots, \mathcal{Z}_N)$*, where* $\mathcal{Z}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ *be a power erosion operator, which is the dual of* $S$*. Then* $S \odot Z$ *is a power relation opening operator.*

The closing is a set dilation operator followed by its dual set erosion operator. Extending this to power relation, a power relation closing operator is a power dilation operator followed by its dual power erosion operator.

**Theorem 10.** *Let* $S = (\mathcal{S}_1, \ldots, \mathcal{S}_N)$*, where* $\mathcal{S}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ *be a power dilation operator. Let* $Z = (\mathcal{Z}_1, \ldots, \mathcal{Z}_N)$*, where* $\mathcal{Z}_n : \mathcal{P}(L_n) \to \mathcal{P}(L_n)$ *be a power erosion operator, which is the dual of* $S$*. Then* $Z \odot S$ *is a power relation closing operator.*

**Definition 20. (Indexed Power Relation)** *A set operator* $\mathcal{S} : \mathcal{P}(L) \to \mathcal{P}(L)$ *maps a subset to a subset. For any* $I \subset [N]$*, an indexed relation with respect to the range sets* $\{L_n\}_{n=1}^N$ *is a pair* $(I, R)$ *where* $R \subseteq \bigtimes_{i \in I} L_i$*.*

*An indexed power relation with respect to the range set* $\{\mathcal{P}(L_n)\}_{n=1}^N$ *is a pair* $(I, \mathbb{R})$ *where* $\mathbb{R} \subseteq \bigtimes_{i \in I} \mathcal{P}(L_i)$*.*

## 5 Future Work

We extend our set operators to multidimensional data by developing the relation theory and the power relation theory. We described the basic definitions, properties and theorems of indexed relation, power relation.

The theory presented has applications in social networks. Centers of activity can be determined by a closing with a small closing operator followed by a larger opening operator followed by a yet large closing operator followed by a yet large opening operator. What will be preserved are centers consisting of nodes each having high interactivity with nodes nearby and what will be cut will be edges whose nodes have low activity with one another.

More advanced kinds of social networks can also be created, where each node instead of having one person or institution has a pair or triplets. Our future work will explore the determination of centers of activity using the opening and closing operators on indexed relations.

## References


1. Xiaojin Ye, Haralick M Robert. The Generalization of Mathematical Morphology to Non-numeric Sets. Intern. Journal Transaction on Machine Learning and Data Mining (October 2021), ISSN (Print) 1865-6781, ISSN (Online) 2509-9337
2. Xiaojin Ye, Haralick M Robert. The Generalization of Mathematical Morphology to Non-numeric Sets. Proceedings, Machine Learning and Data Mining in Pattern Recognition, Petra Perner (Ed.), 17th International Conference on Machine Learning and Data Mining, MLDM 2021, New York, USA, July 17-22, 2021, ibai-publishing, ISSN (Print) 1864-9734, ISSN (Online) 2699-5220

3. Xiaojin Ye. Set Operators.2020
4. Georges Matheron. Random sets and integral geometry. 1975
5. Jean Serra.Image analysis and mathematical morphology. Academic Press, Inc., 1983
6. Banon GJ, Barrera J. Decomposition of mappings between complete lattices by mathematical morphology, part i. general lattices. Signal Processing, 30(3):299-327,1993
7. Banon GJ, Barrera J. Set mapping decomposition by mathematical morphology. chapter to be included in a book entitled "Mathematical Morphology: Theory and Hardware," to be edited by RM Haralick. 1990.
8. Stanley R Sternberg. Grayscale morphology. Computer vision, graphics,and image processing, 35(3):333–355, 1986
9. Henk JAM Heijmans. Morphological image operators, volume 4. Academic Press Boston, 1994
10. A Haas, G Matheron, and J Serra. Morphologie mathematique et granuloetries en place. InAnnales des mines, volume 11, pages 7–3, 1967.
11. Robert M Haralick and Linda G Shapiro. Computer and robot vision, volume 1. Addison-wesley Reading, 1992
12. Edward R Dougherty and Charles Robert Giardina.Image processing: continuous to discrete, volume 1. Prentice-Hall Englewood Cliffs, NJ, 1987
13. Charles R Giardina and Edward R Dougherty.Morphological methods inimage and signal processing. Prentice-Hall, Inc., 1988
14. Robert M Haralick, Stanley R Sternberg, and Xinhua Zhuang. Image analysis using mathematical morphology.IEEE transactions on pattern analysisand machine intelligence, (4):532–550, 1987
15. Petros A Maragos. A unified theory of translation-invariant systems with applications to morphological analysis and coding of images. PhD thesis, Georgia Institute of Technology, 1985
16. Petros Maragos and RW Schafer. Morphological skeleton representation and coding of binary images.IEEE Transactions on Acoustics, Speech, and Signal Processing, 34(5):1228–1244, 1986
17. Petros Maragos. Tutorial on advances in morphological image processing and analysis.Optical engineering, 26(7):267623, 1987
18. Petros Maragos. Pattern spectrum and multiscale shape representation.IEEE Transactions on pattern analysis and machine intelligence, 11(7):701–716,1989
19. Petros Maragos. A representation theory for morphological image and signal processing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(6):586–599, 1989
20. Petros Maragos and Ronald Schafer. Morphological filters–part i: Their set-theoretic analysis and relations to linear shift-invariant filters.IEEE Transactions on Acoustics, Speech, and Signal Processing, 35(8):1153–1169, 1987
21. John L Pfaltz and Josef Šlapal. Transformations of discrete closure systems. Acta Mathematica Hungarica, 138(4):386–405, 2013
22. John L Pfaltz. A category of discrete partially ordered sets. In Mid-AtlanticAlgebra Conf., George Mason Univ., Fairfax VA, 2004
23. John L Pfaltz and Robert E Jamison. Closure systems and their structure.Information Sciences, 139(3-4):275–286, 2001
24. John L Pfaltz. Closure lattices.Discrete Mathematics, 154(1-3):217–236,1996
25. John L Pfaltz and James C French. Multiple inheritance and the closure of set operators in class hierarchies. In INSTITUTE FOR PARALLEL COMPUTATION, UNIV. OF VIRGINIA, 1992
26. John L Pfaltz and John E Karro. Transformations of antimatroid closure spaces. In In 25th Combinatoric, Computing, and Graph Theory Conf.,Boca, 1998

27. Gerhard Gierz, Karl Heinrich Hofmann, Klaus Keimel, Jimmie D Lawson, Michael Mislove, and Dana S Scott.Continuous lattices and domains, volume 93. Cambridge university press, 2003
28. Gabriele Castellini, Jürgen Koslowski, and George E Strecker. Closure operators and polarities. Annals of the New York Academy of Sciences, 704(1):38–52, 1993
29. G Castellini and D Hajek. Closure operators and connectedness. Topology and its Applications, 55(1):29–45, 1994
30. Wikipedia contributors. Closure operator — Wikipedia, the free encyclope-dia, 2017. [Online; accessed 31-August-2018].
31. Marcel Erné, Jürgen Koslowski, Austin Melton, and George E Strecker. Aprimer on galois connections. Annals of the New York Academy of Sciences,704(1):103–125, 1993
32. Relation decomposition: the theory. Robert M Haralick, Ligon Liu, Evan Misshula.International Workshop on Machine Learning and Data Mining in Pattern Recognition.Springer, pages 311–324,2013
33. Xiuyan Ni, Ligon Liu, Robert M Haralick. Music Generation with Relation Join. International Symposium on Computer Music Multidisciplinary Research. Springer. pages 41–64,2016

# The Quest for Artificial Intelligence: A Human Intelligence Perspective

Warwick Graco and Ed lewis

Analytics Shed Jindera NSW Australia 2642
Warwick.graco@analyticsshed.com

**Abstract.**
This paper addresses how what is known about human intelligence (HI) can assist with enhancing the capabilities of artificial intelligence (AI). It addresses the issue of narrow versus broad AI, the brittleness of expert systems and those that use deep learning, the need for hybrid AI which combines deep learning and symbolic reasoning models with incremental and reinforcement learning. They could eventually produce AI systems that display human intelligence. So far AI has fallen well short of emulating what human beings can do when it comes to the ability to comprehend experience and explain what is going on and why. Research into HI has found a factor called '*g*' which is seen to be an ability to educe or abstract patterns and concepts in what is experienced. It is suggested that this eduction ability is integral to other cognitive functions performed by human beings in that it assists them to reason, judge, and decide and that it relies on controlled cognition, such as reflective engagement, self-regulation, and long-term planning; and spontaneous cognition such as mind wandering, daydreaming and intuition. In terms of future research, it is recommended that there is a need to understand more clearly and more conclusively how '*g*' assists people to understand issues and to solve problems and how to develop '*g*' capabilities to assist in the move to develop broad, or strong, AI. Suggestions are made about how to develop AI with HI and examples are provided where it can be applied.

**Keywords:** artificial intelligence, human intelligence, information processing.

## 1 Introduction

This paper addresses what is seen as an 'elephant in the room' to use a cliché and that is the question of what constitutes 'intelligence'? This is that artificial intelligence (AI) has tended to ignore what has been discovered scientifically about human intelligence (HI) [1]. It is suggested that AI will not make significant progress until it incorporates what is known about HI in its capabilities.

AI also called 'machine intelligence' or alternatively 'computational intelligence' is about giving computers and computer-controlled robots the abilities to do tasks that require HI. This includes the abilities of human beings to understand and extract meaning from experience, assists them to exercise will and agency and enables them to do reasoning and problem solving, and learn from mistakes.

HI is studied by psychologists, psychometricians and other scientists who are interested is this issue. Psychometricians have been giving HI attention for well over a century. More broadly, psychometricians study how to measure various psychological attributes such as abilities, traits, attitudes, interests, motives and needs, perceptions, and other types of behaviors [2-4]. They use multivariate statistical techniques, such as factor analysis [5-7], to understand what is measured by tests of mental ability. They have focused especially on the concept of general mental ability or what is referred to by psychometricians and the broader psychological community as '*g'* [8]. They have discovered some interesting insights into '*g*'. A few of these will be explained shortly.

The remainder of this paper will cover briefly key developments in AI and list human capabilities that AI cannot currently perform. This will be followed by what psychometricians have learned so far about HI and how this can assist with achieving comparable capabilities within AI. It will highlight some of the questions and issues that remain to be resolved with understanding intelligence from both a HI and AI perspective. It briefly covers some of the requirements with developing AI with HI. Finally, it provides a few examples of how AI having enhanced human capabilities can provide significant benefits to individuals and the community.

## 2 Key Developments in AI

### 2.1 Narrow versus Broad AI

AI can be divided into categories of narrow, or weak, AI versus broad, or strong, AI [9]. Narrow AI refer to intelligent systems that have learned to carry out specific tasks without being explicitly programmed to do this. There are many applications for narrow AI such as those that assists physicians to diagnose disease, those employed to detect fraud and abuse and those that assist with obtaining the best performance from cars.

Strong AI is a broader and more adaptable form of intelligence that applies to the intricate and skillful tasks such as cutting hair, coming up with a winning design for an advanced fighter jet based on innovations in technology and the performance and logistic requirements the jet is expected to meet, reasoning about what tactics to employ to defeat an opponent in a game of football. This form of AI is starting to emerge, but it has a long way to go to reach advanced stage of development such as having an intelligent personal assistant who can provide sound advice and guidance on complex legal issues and what combination of treatments will help effectively manage a patient who has both covid and diabetes.

### 2.2 Expert Systems

One has been the brittleness of machine-cognition solutions such as expert systems [10,11]. They have not been able to cope with the complexities of the issues that they are dealing with and the exceptions that arise. There are many nuances to an issue, and it is a knowledge engineering challenge trying to accommodate all the subtleties of the problem in an expert system.

Human beings are more able to cope with these issues because of both their abilities to reason and problem solve and the knowledge and experience they have accumulated over time. AI systems are not blessed to the same extent with these capabilities – at least not without a considerable effort on the part of knowledge scientists and knowledge engineers to acquire all the relevant knowledge that applies to the topic or issue under consideration.

Dreyfus [12-14] is another who has been critical of AI and its philosophical foundations. Essentially, he argues that HI and human expertise depend primarily on unconscious processes rather than conscious symbolic manipulation and that these unconscious skills can never be fully included in formal rules.

### 2.3 Deep Learning

Current deep-learning systems [15-17] also are also brittle [18-20]. Deep learning refers to multilayer learning where higher-level features are extracted from the raw inputs. For example, in image processing, lower learning layers may identify edges, while higher learning layers may identify the concepts such as an image of a square.

The brittleness of deep learning systems [18-20] is largely due to machine learning models being based on the 'independent and identically distributed' ('i.i.d.') assumption, which supposes that real-world data has the same distribution as the training data. The 'i.i.d' assumption also assumes that observations do not affect each other (e.g., coin tosses are independent of each other).

This is not a realistic assumption. The data changes for many reasons such as modifications to government policies, the physical movements of people and shifts in weather. Hence AI models and agents must continually adjust and adapt to account for the dynamic circumstances they face.

Bengio, LeCun and Hinton wrote in their paper [20] that the goal of AI is to replicate the kind of general intelligence human beings have. Human beings do not suffer from the problems of current deep learning systems. They and animals can learn massive amounts of background knowledge about the world, largely by observation, in a task-independent manner. This knowledge underpins common sense and allows humans to learn complex tasks, such as driving, with a few hours of practice.

Human beings [18-20] can generalize in a way that is different and more powerful than ordinary 'i.i.d.' generalizations. Human beings can correctly interpret novel combinations of concepts, even if those combinations are unlikely in real life.

### 2.4 Symbolic Systems

Marcus [19] is critical of the views of Bengio, LeCun and Hinton [20-21] and their advocacy of a connectionist deep-learning solutions using neural networks. Marcus argues the merits of a hybrid solution combining deep learning with classical symbolic systems. Symbol manipulation is a very important part of humans' ability to reason about the world. It is also one of the great challenges for deep learning systems.

For example, deep-learning systems rely on images, with little or no comprehension of all the written material that might describe what is evident in the images. In the case of an X-ray, this could include a patient's medical history, and this could result in critical information being missed.

The deep learning advocates [20] maintain their conviction that better neural network architectures will eventually incorporate all aspects of human and animal intelligence, including symbol manipulation, causal inference, and common sense.

### 2.5 Hybrid Systems

Hybrid solutions [19,22] combine the connectionist neural net approach because of its strength with pattern recognition with symbol manipulation because of its strengths of providing context to problems and issues and explaining why an object recognized by a neural network is, for example, a cancerous cell. That is, connectionist solutions will provide the 'what' (i.e., results) while symbolic ones will provide the 'why' (e.g., why cases are classified high risk). Of course, symbolic solutions can also provide the 'what' as well with, for example, identifying cases that are high risk based on the results of applying business rules.

### 2.6 Inductive Logic Programming

An example of a robust approach to symbolic reasoning is inductive logic programming (ILP) [23]. This is the subfield of machine learning that uses first-order logic to represent hypotheses and data.

An example of inductive reasoning is that I had coffee once at the cafe and it was horrible, so all its coffee must be awful. Inductive reasoning begins with specific observations and measures of issues under consideration and then identifies patterns and regularities in the data, formulates tentative hypotheses that can be tested with the aim to come up with conclusions that follow from the facts.

ILP is differentiated from most other forms of machine learning both by its use of an expressive representation language and its ability to make use of logically encoded background knowledge.

ILP techniques can be used for purposes such as learn the structure-activity rules for drug design, develop fault diagnosis rules for satellites, generate viable and testable hypotheses from ecological data such as 'who eats who' in a habitat, to working out rules for playing games such as checkers.

ILP has many appealing features including that it can generalize from small numbers of examples, it can analyze complicated issues to represent facts in symbolic form, and it can explain the reasons for its conclusions.

ILP has the challenge that it is hard to use and that therefore can require a skilled developer to build an ILP system. It also requires human experts to provide facts on the issue being analyzed.

## 2.7 Incremental Learning

First order logic systems can be enhanced using incremental learning [24]. This is a method of machine learning in which input data is used continuously to extend the existing model's knowledge. It can be applied when training data becomes available gradually over time or when system memory limits are exceeded. The aim of incremental learning is for the learning model to adapt to new data without disregarding what it previously learned. Manually building models can be a complex, costly, and error-prone practice. Incremental development of models, and the ability to learn complex conditions that apply to tasks being learned are also desirable.

## 2.8 Reinforcement Solutions

Nicholson [25] stated that deep reinforcement learning, where machines learn by testing the consequences of their actions and receiving feedback, is one of the most promising areas of AI. It combines deep learning with reinforcement learning which together can be trained to achieve goals. It's a crucial part of the learning for self-driving vehicles and industrial robots, which must navigate complex environments safely and on time.

Reinforcement learning can do much more than control individual machines. It can direct entire orchestras of machines, steer complex systems toward better performance, route fleets, and coordinate robotic teams.

Another advantage with reinforcement learning is that simulations can be used to train and reinforce patterns of behavior for situations that occur rarely, if not at all, such as in emergencies. This way systems are more robust and capable when the unexpected occurs.

Dickson [26] stated that reinforcement learning can be employed to develop abilities associated with intelligence such as perception, motor functions and language. This is in contrast with AI systems that try to replicate specific functions of natural intelligence such as classifying images, navigating physical environments, or completing sentences.

AI researchers go as far as suggesting that with well-defined rewards, a complex environment, and the right reinforcement learning algorithm, that it will be possible to reach artificial general intelligence or strong AI. This is the kind of problem-solving and cognitive abilities found in humans and, to a lesser degree, in animals.

The use of reinforcement learning agents; that can outperform human beings in games such as Go, chess, Atari, StarCraft, and other games; have been developed [19]. They have also developed reinforcement learning models to make progress in some of the most challenging issues of science such as solving intractable mathematical problems [27].

## 2.9 Stock Take

This brief overview of AI developments indicates that early attempts to develop intelligent solutions have not been successful but more recent ones that combine connectionist systems and symbolic models combined with incremental and reinforcement learning look promising. They could eventually produce AI systems that display human intelligence.

The challenge remains to emulate what human beings can do when it comes to HI and that is learn massive amounts of background knowledge about the world, largely by observation, in a task-independent manner. They can learn complex tasks such as operating a computer. They can discriminate what is important in what is being learned and they can explain why things do not work. They can use their powers of reasoning to solve problems.

AI falls well short of showing these capabilities at present. The basic problem is that it can perform tasks such as high-speed arithmetic, data processing, pattern recognition and even judgment, decision making and problem solving. It can do all these tasks in a mechanical manner either by programming or by learning from examples but show little or no evidence of intelligence. That is, the ability to comprehend experience and explain what is going on and why.

## 3 Research into Human Mental Abilities

This is a segue to what has been found about the intelligence of human beings. In a nutshell it has gone from (1) the discovery of intelligence being a single ability, (2) to attempts to show that it consisted of several abilities (3) to a view that intelligence is a hierarchy with a single overarching ability at the top and a series of aptitudes in the middle and very specific abilities at the bottom in what is a pecking order of abilities. Recent approaches have tended to give attention to processing models of HI. For those who want further information on these developments they can consult Kaufman et al [28], Neisser et al [29] and Sternberg [30].

These developments in the gaining of understanding of HI are not discussed further in this paper and instead two issues are focused on below. The 1st is the existence of a single overarching ability with HI and the 2nd is to draw attention to the information-processing aspects of intelligence. This is done because firstly, despite attempts by some to 'dethrone' the theory of a single ability of HI, it has been an issue that refuses to go away and keeps resurfacing. For an example of this trend see Canivez and Youngstrom [31]. Hence attention is focused on this single ability in the discussion below. It will also be argued below that for AI to migrate from narrow, or weak, AI to broad, or strong, AI one way this can be done is by understanding the steps employed to solve problems and to respond to issues and the mental and other functions involved.

Turning to the issue of a single ability of HI, the 1st significant breakthrough in the analysis of human mental abilities was the discovery of the '*g*' factor that was highlighted at the beginning of this paper. This construct was discovered in investigations of mental abilities by Spearman [32,33] in the early years of the 20th Century. Spearman found it using factor analysis [5-7]. This is a data reduction technique that discovers the underlying constructs called 'factors' that explain the correlations between, for example, tests of mental ability. Spearman saw '*g*' as being a single overarching mental ability that explains performance on tasks requiring understanding and solving problems.

### 3.1 Eduction Ability

Spearman [32,33] formed the view that what is measured by '*g*' is what he called the 'education of relations' and the 'education of correlates'. One interpretation of what Spearman was inferring with education is that '*g*' is a measure of a person's ability to make sense of problems and issues. That is, before a problem can be solved, the problem solver must perceive and comprehend the problem that must be addressed. This includes the ability to 'educe' or abstract the concepts or rules that enable the problem to be solved.

A simple example is working out the next number in the series (1 2 4 8)? The problem solver must educe the rule that each number is being doubled to provide the next number. So, the next number in the above series is 2*8 which is 16.

It would be misleading to assert that this education ability is all that there is to HI. To the contrary, it can be seen as an integral ability that enables other cognitive functions to be performed such as deductive and inductive reasoning, problem solving, reaching decisions, and making judgments. In other words, HI is a range of integrated abilities involving recalling information, grasping relationships, spotting differences and similarities, calculating results and drawing inferences from what is known. Eduction is critical to the success of these functions as it enables the individual concerned to comprehend experience and draw meaning from what is processed.

### 3.2 Information Processing

Another way of construing this process is that people go through a series of steps to achieve goals, to solve problems and to respond to issues. These steps include (1) perceiving the situation, (2) making sense of it, (3) reasoning about it such as doing inductive and deductive thinking to gain greater clarity and understanding of what is being contemplated (4) making judgments of the worth of possible solutions and (5) and making a choice about which option is the best one to enable the outcome to be achieved.

People perform information-processing steps at both conscious and unconscious levels. Human beings make numerous decisions each day such as when to get out of bed, what food to eat and when to press the brake when driving a car. Many decisions and actions are automatic in that people do them habitually while others are deliberate, conscious ones [34].

Kaufman et al [28] and Kaufman [35] conceptualize this distinction between deliberate and automated procedures/responses as one of controlled (or goal-directed) versus spontaneous cognition. Kaufman et al [28] discerned a number of differences between these two modes of cognition in their dual process theory.

The 1st was that controlled processes consume limited attentional resources. In contrast, spontaneous ones are not dependent on inputs from these higher-level control processes.

The 2nd was they are not fixed or static processes. They are constantly changing through learning and experience as people engage continually with the world. Individuals develop skills because of what they learn and experience.

The 3rd is that controlled cognition involves both the ability and the inclination across situations to think about thinking. This is called 'metacognition' [36]. This covers the abilities to reflect on prior experience and behavior and use the insights gained to modify thinking and habits as examples.

The 4th is that the mental and other functions that are part of the controlled cognition include reflective engagement, self-regulation, self-control, perseverance, long-term planning, dissociable components of executive functioning such as working memory, cognitive and affective inhibition, and mental flexibility, explicit cognitive ability or the skill set that lies at the heart of highly *g*-loaded tasks, intellectual engagement, and elementary cognitive tasks that support explicit cognitive ability. What links these processes together is that they all draw on the limited attentional resources of people.

The 5th is that the individual differences in spontaneous cognition reflect the ability to acquire information automatically and the tendency to engage in unprompted forms of cognition. For instance, whereas most people can spontaneously experience gut feelings and daydreams, there may be individual differences in the extent to which people are willing to engage with them. Mental and other functions that are part of the spontaneous cognition include mind wandering, daydreaming, implicit learning, latent inhibition, intuition, acquired forms of expertise and long-term memory, and implicit domains of mind that are universal human domains pertaining to knowledge of spatial relations, number, probability, logic, language, people, language, music, aesthetics, living things, the inanimate physical world, or the beliefs and desires of other minds.

### 3.3 Some Key Challenges

Kaufman et al [28] stated that one study found that individual differences in implicit learning predict intelligent behaviors such as language learning and verbal analogical reasoning above and beyond '*g*' and the cognitive mechanisms underlying '*g*'. This highlights a few challenges for those involved in AI and HI going forward. They include:

- The need to understand more clearly and more conclusively how '*g*' assists people to understand issues and to solve problems
- How to develop '*g*' capabilities with AI to hopefully assist with developing broad or strong AI capabilities

They are both difficult and demanding challenges and could require decades of more research and development to achieve AI with HI. The 1st task includes understanding how '*g*' assists with both controlled and spontaneous cognition.

### 3.4 Another Challenge

Another challenge that requires further research is what role does '*g*' play in the performance of complex and demanding tasks such as performing surgery, flying an aircraft, or managing a corporation. For example, Horn and McArdle [37] stated that the reasoning that is regarded as indicating intelligence in chief executive officers of organizations is expertise deductive reasoning [38]. It is the kind of reasoning identified in descriptions of the thinking of experts in chess, financial planning, and medical diagnosis [39-41]. Horn and McArdle also stated that wide span working memory [42] that also appeared to be important with executive management performance. Both capabilities were judged by Horn and McArdle [37] to be indicators of HI but were found not to be related to existing mental abilities that underlie HI. The question is, if this is the case, is there a need to broaden the abilities that are seen to constitute HI?

## 4 Developing AI with HI Capabilities

This is a speculative issue to address because it means trying to forecast future developments when there are many unknowns and uncertainties. However, there are clues about what will lead to the breakthrough of giving AI the capabilities of human beings.

What are the clues? These were covered previously in the paper. A core requirement will be exploiting the strengths of deep learning and the symbolic reasoning systems combined with incremental and reinforcement learning to abstract or educe, label and store patterns and concepts found in data and use these to facilitate sense making, problem solving and decision making.

There is also need to compare what is discerned to what is known, reason about it, store what is learned in memory and apply what is learned to new cases. It is also important to overcome the limitations of machine learning where they require many examples to learn patterns and concepts in data.

Another requirement is to assemble and store a massive amount of background knowledge as this underpins common sense and allows humans to learn complex tasks, such as driving, with a few hours of practice. This knowledge also assists with reasoning and learning.

In practice this will mean using deep learning to do pattern recognition and feature selection. These steps are examples of education in that they involve abstracting key concepts in examples being processed. This abstraction process will be aided by the use the symbolic system to reason about what is found in the examples and the encyclopedic knowledge stored in memory to assist with understanding what is educed. For example, what is stored in memory can be compared to what is abstracted to work out that what is perceived is an image of a giraffe.

There is the corresponding need to emulate what human beings excel at and that is learning from one or more examples. It will require using deep learning, the symbolic reasoning system and what is stored in memory to perform analysis and synthesis to understand better what is abstracted from each example and how it can be applied.

The next step is to determine how what is perceived and processed is managed or treated. For example, if it is a self-driving intelligent vehicle what does it do if there is a large dog standing in the middle of the road 100 meters in front of the vehicle?

This leads to the next issue that the steps outlined above dovetail in with the information processing architecture put forward by LeCun [43]. This includes:

- The **perception module** that receives signals from sensors and estimates the current state of the world.
- The **world model module** that constitutes the most complex piece of the architecture. Its role is twofold: (1) to estimate missing information about the state of the world not provided by perception, and (2) to predict plausible future states of the world.
- The **cost module** that computes a single output that predicts the level of discomfort of the agent. It is composed of two submodules: (1) the intrinsic cost one which is hard-wired and immutable (not trainable) and computes the immediate discomfort (such as damage to the agent, violation of hard-coded behavioral constraints, and similar), and (2) the critic which is a trainable submodule that predicts future values of the intrinsic cost. The goal of the agent is to minimize the intrinsic cost over the long run.
- The **actor module** that computes proposals for action sequences. The actor can find an optimal action sequence that minimizes the estimated future cost, and output the first action in the optimal sequence, in a fashion like classical optimal control.
- The **short-term memory** module that keeps track of the current and predicted world state, as well as associated costs.
- The **configurator module** that performs executive control: Given a task to be executed, it preconfigures the perception module, the world model, the cost, and the actor for the task at hand, possibly by modulating the parameters of those modules.

This solution does not accommodate the ability to abstract concepts and rules in what is perceived and a symbolic reasoning system. It is suggested if these were incorporated in this architecture it would provide an information-processing system that can help drive the future development of AI. That is, moving it to having the capabilities of human beings or alternatively, moving from narrow to broad or strong AI.

## 5 Some Important Applications

AI is ubiquitous in that it is applies to all areas of the community such as government, transport, manufacturing, commerce, health, education, sports, entertainment, and agriculture to give examples.

There are three technologies where AI that possesses HI are highly likely to confer significant benefits on the community. The first and obvious technology is smart robots which comprehend, think, and respond to human commands. These will include droids like those seen in the Star Wars Film series and cyborgs found in many science-fiction books and films.

For example, smart robots will be able to assist the elderly who live at home and lack the strength and mobility to do domestic tasks. This will take pressure off the use of nursing homes [44] that are currently in crisis because of the covid pandemic and reduce the need for governments to provide these facilities for those too old to look after themselves.

It is estimated [45] that there will be a shortage of between 37,000 and 124,000 physicians in the USA by 2034. Smart robots in the form of robot doctors [46] and robot nurses [47] can be used to cover for these projected shortages in health professionals.

The second is developing smart self-driving vehicles and assisting them to identify other vehicles and users of the roads, comply with road rules and take passengers safely to their destinations. Clarke [48] reports that those developing these vehicles are stuck on the last 10 percent of the developmental curve where these vehicles must respond to rare and unusual events such as a ball bouncing across the street followed by a running child. The problem is complicated by current self-driving vehicles acting supremely confidently when they are wrong such as turning on a railway crossing to follow a railway line. It is suggested that smart vehicles with HI capabilities will help overcome these difficult challenges.

Chatbots [49] are virtual assistants that interact with people using either written or spoken words usually in online conversations. They answer simple queries and are used in business-to-consumer and business-to-business transactions with clients.

The next stage in the development of these agents will be smart digital assistants. These too will be intelligent agents with human capabilities. They will perform secretarial tasks for their 'masters' such as arrange appointments, organize travel and accommodation, order food and beverages, and assist them with sense making, problem solving and decision making. They will do translations of documents not written in languages their 'masters' understand, search and find information their 'masters' require to perform tasks, do analyses of data, and perform complex calculations, and provide advice and guidance. They will be essentially peoples alter egos and will assist them with difficult decisions such as which car to purchase, what insurance policy to take out, where to live and what accommodation is best suited to their needs and indicate the friends and partners, they are likely to have the greatest rapport with and gain the greatest satisfaction from when it comes to interpersonal relationships. They will also monitor the health and fitness of their 'masters' and advise when they need to see a health professional.

The most important advantage with this technology is to help prevent their 'masters' from making poor and costly decisions such as making bad investments, marrying the wrong partner, and selecting an inferior career and applying for less-than-ideal jobs. This has the potential to reduce economic, legal, welfare, health and educational costs that arise from making suboptimal decisions. This is illustrated by Till [50] who found in a survey that about one in five people are in the wrong job and that moving them to a role they like would result in an estimated 3.6 percent increase in revenue or about $36m for company earning a billion dollars a year.

## 6 Conclusion

To summarize, this paper covered the key developments in AI and how these can be further enhanced to give AI the important cognitive capabilities of human beings. This includes using the strengths of deep learning, symbolic reasoning system, incremental learning, reinforcement learning, and the knowledge stored in memory to assist AI to think and act in an intelligent manner. It was also suggested that these outcomes will be accomplished by developing AI with information processing capabilities.

In conclusion, it is considered that ‘*g*’ is what is missing from AI and that it will not make significant progress until this issue is given serious attention by AI researchers. This is seen as a key challenge going forward with this area of research.

## References

1. Gottfredson, L. S.: Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. Intelligence 24, 13–23 (1997)
2. Furr, R.M.: Psychometrics: An Introduction. Sage Publications, Thousand Oaks CA (2021)
3. Rust, J; Kosinski, M; Stillwell, D.: Modern Psychometrics: The Science of Psychological Assessment. Routledge, Milton Park, Abingdon-on-Thames, Oxford Shire UK (2020)
4. Price, L.R.: Psychometric Methods: Theory into Practice. Guilford Press, New York NY (2016)
5. Gorsuch, R.L.: Factor Analysis. Routledge, Milton Park, Abingdon-on-Thames, Oxford Shire UK (2015)
6. Thompson, B.: Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications. American Psychological Association, Washington DC USA (2004)
7. Kline, P.: An Easy Guide to Factor Analysis. Routledge, Milton Park, Abingdon-on-Thames, Oxford Shire UK (2014)
8. Gottfredson, L. S.: The General Intelligence Factor. Scientific American 4, 24-30 (1998)
9. Heath, N.: What is AI? Here’s everything you need to know about artificial intelligence? ZDNET Special Feature: Managing AI and ML in the Enterprise (2021) https://www.zdnet.com/article/what-is-ai-heres-everything-you-need-to-know-about-artificial-intelligence/
10. Kumar, A.V.S; Kalpana, M.: (eds.) Fuzzy Expert Systems and Applications in Agricultural Diagnosis (Advances in Environmental Engineering and Green Technologies) IGI Global, Hershey Pennsylvania (2019)
11. De Brito, J; Pereira, C; Silvestre, J.D; Flores-Colen, I.: Expert Knowledge-based Inspection Systems: Inspection, Diagnosis, and Repair of the Building Envelope. Springer, Berlin Germany (2020)
12. Dreyfus, H.: Alchemy and Artificial Intelligence. RAND Corporation, Santa Monica CA (1965)
13. Dreyfus, H.: What Computers Still Can’t Do. MIT Press, New York NY (1992)
14. Dreyfus, H; Dreyfus, S.: Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. Blackwell, Oxford UK (1986)
15. Aggarwal, C.C.: Neural Networks and Deep Learning: A Textbook. Springer, Berlin Germany (2018)
16. Graupe, D.: Deep Learning Neural Networks. World Scientific, Singapore (2016)
17. Patterson, J; Gibson, A.: Deep Learning: A Practitioner’s Approach. O’Reilly Books, Sebastopol California USA (2017)
18. Dickson, B.: The future of deep learning, according to its pioneers. TechTalks Newsletter 1 Jul (2021) https://bdtechtalks.com/2021/07/01/deep-learning-future-bengio-hinton-lecun/
19. Marcus, G.: Deep Learning is Hitting a Wall. Medium 10 March (2022) https://nautil.us/deep-learning-is-hitting-a-wall-14467/?utm_medium=email&utm_source=topic+optin&utm_campaign=awareness&utm_content=20220328+data+ai+nl&mkt_tok=MTA3LUZNUy0wNzAAAAGDbTl-

suEtfbDPvKFFwC-Py5yuCBikrdlju6c5WYygauMPIxyR_8KxFvXnqzm9QDB-qYpAMApnVIo-qjXwhMR9qBe-ekiTJYR8Y9j-5zGF0OGQp6RG
20. Bengio,Y; LeCun, Y; Hinton, G.: Deep Learning for AI. Communications of the ACM 64 (7), 58-65 (2021)
21. LeCun, Y; Bengio, Y; Hinton, G.: Deep Learning. Nature 521 (7553), 436-444 (2015)
22. Dickson, B.: The Case for hybrid artificial intelligence. TechTalks Newsletter 4 March (2020) https://bdtechtalks.com/2020/03/04/gary-marcus-hybrid-ai/
23. Cropper, A; Dumanci, S; Muggleton, S.H.: Inductive logic programming at 30. Machine Learning 111,147–172 (2022) https://doi.org/10.1007/s10994-021-06089-1
24. Luo, Y; Yin, L; Bai, W; Mao, K.: An Appraisal of Incremental Learning Methods. Entropy 2020*,* *22,* 1190. https://doi.org/10.3390/e2211119
25. Nicholson, C.: Deep reinforcement learning helps us master complexity. TechTalks Newsletter 22 July (2021) https://bdtechtalks.com/2021/07/22/deep-reinforcement-learning-complexity/
26. Dickson, B.: Evolution, rewards, and artificial intelligence. TechTalks Newsletter, 17 June (2021) https://bdtechtalks.com/2021/06/17/evolution-rewards-artificial-intelligence/
27. Rodriguez, J.: Inside DeepMind's New Efforts to Use Deep Learning to Advance Mathematics. KDNuggets December (2021) https://www.kdnuggets.com/2021/12/inside-deep-mind-new-efforts-deep-learning-advance-mathematics.html
28. Kaufman, J.C; Kaufman, S.B; Plucker, J.A.: Contemporary theories of intelligence. In J. Reisberg (ed.) Handbook of Cognitive Psychology. Chapter 51. Oxford University Press, New York, NY (2013)
29. Neisser, U; Boodoo, G; Bouchard, T.J; Boykin, A.W; Brody, N; Ceci, S.J; Halpern, D.F, Loehlin, J.C; Perloff, R; Sternberg, R.J; Urbina, S.: Intelligence: Knowns and Unknowns. American Psychologist 51(2), 77-101 (1996)
30. Sternberg, R.J.: (ed.) The Cambridge handbook of Intelligence. Cambridge University Press, New York NY (2020)
31. Canivez,G.L; Youngstrom, E.A.: Challenges to the Cattell-Horn-Carroll Theory: Empirical, Clinical, and Policy Implications. Applied Measurement in Education 32 (3), 232-248 (2019)
32. Spearman, C.: General Intelligence, Objectively Determined and Measured. The American Journal of Psychology 15 (2), 201-292 (1904)
33. Spearman, C. Theory of two factors. Psychological Review 21,101-115 (1914)
34. Kahneman, D.: Thinking, Fast and Slow. Farrar, Straus and Giroux, New York NY (2013)
35. Kaufman, S.B.: Ungifted: Intelligence Redefined. Basic Books, New York NY (2013)
36. Chick, N.: Metacognition. Vanderbilt University Center for Teaching. (2013) https://cft.vanderbilt.edu/guides-sub-pages/metacognition/
37. Horn J.L; McArdle J.J.: Understanding human intelligence since Spearman. In: Cudeck R; MacCallum R. (eds.) Factor analysis at 100 years. pp. 205–247. Erlbaum, Mahwah, NJ (2007)
38. Charness, N.: Expertise in chess: the balance between knowledge and search. In Ericsson K.A; Smith, J. (eds.) Toward a general theory of expertise. pp. 39-63. Cambridge University Press, New York NY (1991)
39. Chi, M. T. H; Glaser, R; Farr, M. J.: (eds.) The nature of expertise. Erlbaum, Hillsdale, NJ (1988)
40. de Groot, A.D.: Thought and choice and chess. Mouton, The Hague (1978)
41. Ericsson, K. A; Pool, R.: Peak: Secrets from the new science of expertise. Houghton Mifflin Harcourt, New York NY (2016)
42. Ericsson, K. A; Kintsch, W.: Long-term working memory. Psychological

Review 105, 211-245 (1995)
43. Yann LeCun on a vision to make AI systems learn and reason like animals and humans. Meta AI 23 February (2022) https://ai.facebook.com/blog/yann-lecun-advances-in-ai-research/
44. Daly, R.: Majority of nursing homes are operating at a loss, survey finds. Healthcare Financial Management Association. 24 August (2020) https://www.hfma.org/topics/news/2020/08/majority-of-nursing-homes-are-operating-at-a-loss--survey-finds.html
45. AAMC Report Reinforces Mounting Physician Shortage. Association of American Medical Colleges. 11 June (2021) https://www.aamc.org/news-insights/press-releases/aamc-report-reinforces-mounting-physician-shortage#:~:text=MEDIA%20CONTACTS&text=According%20to%20new%20data%20published,both%20primary%20and%20specialty%20care
46. Trafton, A.: The (robotic) doctor will see you now. MIT News. 4 march (2021) https://news.mit.edu/2021/robotic-doctor-will-see-you-now-0304
47. Christoforou, E. G., Avgousti, S., Ramdani, N., Novales, C., Panayides, A. S.: The Upcoming Role for Nursing and Assistive Robotics: Opportunities and Challenges Ahead. Frontiers in digital health, 2, 585656 (2020) https://doi.org/10.3389/fdgth.2020.585656
48. Clarke, L.: How self-driving cars got stuck in the slow lane. The Guardian. 28 March (2022) https://www.theguardian.com/technology/2022/mar/27/how-self-driving-cars-got-stuck-in-the-slow-lane?utm_term=6240fa834fb68e017c8ef7f058d4fa10&utm_campaign=GuardianTodayAUS&utm_source=esp&utm_medium=Email&CMP=GTAU_email
49. Selig, J.: The Power of Chatbots Explained. Expert.AI. 24 March (2022) https://www.expert.ai/blog/chatbot/
50. Till, I.: One in five people is in the wrong job. Right Management. (2015) https://www.rightmanagement.com.au/wps/wcm/connect/633e96f8-23de-42b7-8ccb-9512cb9ec202/Candidate+Survey+Whitepaper+-+270815-5.pdf?MOD=AJPERES

# Authors Index