



Deliverable 4.1

FROM MEDICAL DATA TO KNOWLEDGE: STATE OF THE ART ANALYSIS

WP4 – Knowledge Extraction and Data Processing
T4.1 State of the art analysis

Document Type	Document
Document Version	V1.0
Access Level	Public
Contractual Submission Date	February 2024
Actual Submission Date	April 2024
Editors	icomatrix
Contributors	ViraSoft, Bilbest, Mantis

Document Contributors

Partner	Author	Role
icomatrix	Diana Sima	Contributor-Editor
ViraSoft	Sercan Cayir	Contributor
Bilbest	Suheyla Turkyilmaz	Contributor
Mantis	Behzad Naderalvojoud	Contributor-Reviewer

Document History

Date	Version	Editors	Status
05/01/2024	0.0	icomatrix	Draft (Table of Contents)
24/01/2024	0.2	ViraSoft	Draft (Cancer use case)
20/02/2024	0.5	icomatrix	Draft (MS use case)
10/03/2024	0.7	icomatrix	Draft (General)
25/03/2024	0.9	Bilbest & Mantis	Draft (Sepsis use case)
15/04/2024	1.0	Mantis	Final (Review)

Table of Contents

1	Extraction of health knowledge from medical sources	6
1.1	Defining health knowledge extraction problems	6
1.2	Machine learning methodologies for health knowledge extraction problems	6
2	MS Use Case	7
2.1	Robust imaging biomarkers for diagnosing and monitoring multiple sclerosis	7
2.2	Predicting clinical scores and disease progression in MS patients using brain MRI scans	7
2.3	Predicting treatment response in MS	8
2.4	Predicting conversion to definite MS or secondary progressive MS	9
3	Breast Cancer Use Case	10
3.1	Synoptic pathology report examination	10
3.2	Double-check for the pathology outcomes	10
3.3	3-month mortality in metastatic disease	12
4	Sepsis Use Case	13
4.1	Machine learning applications for sepsis	13
4.2	Machine learning techniques used in sepsis applications	14
5	Conclusions	15

Glossary

Acronym	Meaning
AI	Artificial Intelligence
AUROC	Area Under Receiver Operating Characteristic Curve
BERT	Bidirectional Encoder Representations from Transformers
CNN	Convolutional Neural Network
CONCH	CONtrastive learning from Captions for Histopathology
DMT	Disease Modifying Treatment
EBM	Evidence-Based Medicine
EDSS	Expanded Disability Status Scale
EHR	Electronic Health Record
HeKDisco	Healthcare Knowledge Discovery
LSTM	Long Short-Term Memory
MIDOG	MItosis DOnain Generalization
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
RNN	Recurrent Neural Networks
VLM	Visual Language Model
XGBoost	eXtreme Gradient Boosting

EXECUTIVE SUMMARY

The HeKDisco project proposes a novel knowledge discovery process in health care systems that will provide physicians with reliable evidence on various treatment stages and clinical events, thereby reducing individual clinical errors. HeKDisco, following evidence-based medicine (EBM), aims to use the best (reliable) evidence in making decisions about the care of individual patients so that clinician's experience, patient's values and preferences, and the best empirical clinical guidelines are integrated. HeKDisco aims to transfer big health data from volume-based to value-based by generating a relational knowledge base that can lead to innovative treatments, predict therapeutic outcomes, and early diagnosis. The project plans to demonstrate its prototype models in several medical use cases in the fields of neurology/neuroimaging (multiple sclerosis (MS) use case), pathology/oncology (breast cancer use case), infectious diseases (sepsis use case), and elderly care (fall risk use case).

This deliverable describes state-of-the-art solutions for the extraction of health knowledge from medical sources. Relevant approaches are analyzed and selected for use in the project, in alignment with the use cases tackled in HeKDisco. Existing solutions are described, and arguments are made for the choice of methods used in the project. This deliverable is part of WP4, "Knowledge Extraction and Data Processing", which focuses on the definition of model architectures, of input and output data, and on the development of knowledge extraction models for each use case.

1 Extraction of health knowledge from medical sources

Numerous types of methods can be considered to fall under the umbrella of *knowledge extraction from medical sources*. We restrict this report to one of the main types of methods: machine learning. Machine learning models automatically extract salient information for pre-defined tasks, typically through the use of labeled data acting as training examples. They can make use of health-related information from medical sources (inputs) for solving particular tasks (e.g., diagnosis or outcome prediction). Related techniques of data mining and statistics allow for the discovery of patterns and insights from large (medical) datasets. In the medical field as well as other domains, good standardization of data is a prerequisite for applying machine learning. This can be achieved by utilizing ontologies in order to structure and extract health-related information from medical records.

In this section, we enumerate typical problem formulations and types of machine learning algorithms that are applicable to the use cases considered in the HeKDisco, and more generally in the health knowledge field.

In sections 2, 3 and 4, we dive deeper into the specific challenges, problem formulations and state-of-the-art algorithms for three use cases (MS, breast cancer and sepsis). We examine domain-specific methodological advances from the recent literature and provide insights into going beyond the state-of-the-art in this project.

1.1 Defining health knowledge extraction problems

Each medical application requires adequate formulations for each specific knowledge extraction problem that can be addressed with computerized models and infrastructure. Despite variations across domains, there exist common problem formulations that allow a direct link between problem definition and the selection of the most suitable machine learning algorithms to solve each problem. As such, medical knowledge extraction problems can be formulated, in many domains, as either classification, regression, pattern recognition, image segmentation, anomaly detection, etc. This opens a wide range of possibilities for the resolution of each problem, using modern AI paradigms and software.

1.2 Machine learning methodologies for health knowledge extraction problems

Historically, healthcare knowledge discovery relied on addressing specific (hypothesis-driven) problems using relatively small datasets collected from single sources, in a homogeneous manner. Accordingly, simple machine learning approaches, such as univariate models and decision trees, were sufficient. Nowadays, there is a paradigm shift towards large, multi-center, heterogeneous datasets. This requires more complex modeling, which is today technically possible due to advances in hardware (computational power and data storage capabilities) and distributed IT infrastructures (cloud computing, network protocols), offering scalable resources, seamless communication and secure data exchange. Algorithmically, advances in AI that have

revolutionized image processing and natural language processing, are also seen to enter the medical field. Indeed, convolutional neural networks (CNNs) became a cornerstone in processing medical images (Çiçek et al, 2016) and deep learning methods such as the Bidirectional Encoder Representations from Transformers (BERT) have been used for extracting parameters from medical texts. Moreover, deep learning frameworks such as foundation models, and the use of ensemble models, provide novel approaches for dealing with huge, heterogeneous medical datasets, and show important potential in the field of healthcare knowledge extraction.

2 MS Use Case

2.1 Robust imaging biomarkers for diagnosing and monitoring multiple sclerosis

Multiple sclerosis (MS) is a chronic disease that affects almost 3 million people worldwide, with more than 700,000 people in Europe. There are more than 20 disease modifying treatments (DMTs) currently available to slow down, but not cure, the disease. Diagnosing MS involves a combination of clinical evaluation, medical history analysis, and various diagnostic tests. There is no single test that definitively confirms the presence of MS. Magnetic Resonance Imaging (MRI) is a crucial tool in diagnosing MS. It shows the presence of lesions (or plaques) in the brain and spinal cord, providing essential diagnostic information (Filippi et al., 2019). The current MS diagnostic criteria, the McDonald criteria (Thompson et al, 2018), aim to provide objective evidence that the central nervous system is attacked in multiple regions (dissemination in space), and at different points in time (dissemination in time).

Numerous algorithms for MS lesion detection and quantification on brain MRI scans have been proposed in the past decades (see (Danelakis et al., 2018) for a review). Performant automatic MS lesion identification and segmentation software tools use deep learning approaches such as 3D U-net convolutional neural networks (Çiçek et al, 2016). Brain MRI scans can also be employed to estimate brain atrophy, another essential biomarker in MS that provides insights into disease progression and neurodegeneration (Popescu et al., 2013). **icomatrix's** flagship product **icobrain ms**, which is used for processing brain MRI scans in the HeKDisco MS use case, is already equipped with deep learning-based lesion segmentation (Rakić et al, 2021) and robust brain atrophy computations (Smeets et al, 2016).

2.2 Predicting clinical scores and disease progression in MS patients using brain MRI scans

Predicting disease progression in MS in terms of reaching a certain disability stage is of high interest. Apart from brain imaging, progression of MS is often characterized via a series of clinical tests which determine the disability of a patient. The most commonly used and recognised is the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983), which assigns a score from 0 to 10 in 0.5 increments depending on the patient's physical impairment.

Most prognostic studies in MS so far have been examining known biomarkers or new candidate biomarkers using univariate models or classical linear multivariate statistics, taking widely accepted clinical scores such as EDSS as outcome measures. Different individual MRI biomarkers have been used as predictors of disability progression. These included measures of lesion volume, whole brain atrophy, and specific regional atrophy patterns. *icometrix*' lesion and brain volume measurements were already found useful in explaining part of the variance in EDSS and cognitive scores in a real-world setting, with a dominating role for whole brain volume in the model of physical disability and for lesion volume in the model of cognitive disability (D'hooghe et al, 2019).

Directly predicting the EDSS score or assessing the time needed for reaching a certain EDSS disability milestone have been attempted. The first end-to-end deep learning model using brain MRI for predicting disability progression in terms of EDSS one year from baseline achieved an area-under-the-curve of only 0.66 when using raw baseline brain MRIs as input, which increased to 0.70 when including annotations of MS lesion as input as well (Tousignant et al, 2019). Some recent studies have obtained more promising results for future prediction of disability (Roca et al., 2020, Storelli et al., 2022, Coll et al, 2023) and cross-sectional patient stratification (Cruciani et al., 2021, Marzullo et al., 2019).

Prediction models are however hampered by the fact that they might learn an average trend from the data, without sufficiently taking into account the heterogeneity of MS. A more interesting direction has been described by Eshaghi et al (2021), who focused on discovering temporal and phenotypic characteristics in the MS brain. They used MRIs collected in several clinical trials (placebo arms) and selected only 13 features derived from brain MRI scans (such as regional brain volumes) that were found to be most significantly different between MS and healthy controls; then they modeled the evolution of these features in untreated MS patients with a Subtype and Staging Inference model (Young et al, 2018). This experiment led to the identification of three subtypes, with distinct patterns of evolution, which were interpreted on the basis of the earliest MRI abnormality as cortex-led, normal-appearing white matter-led, and lesion-led. These subtypes, in turn, were shown to be correlated with disability progression in untreated patients, as well as with how patients responded to therapy (as measured on the EDSS scale).

2.3 Predicting treatment response in MS

With the advent of DMTs, the question is also how to select patients with high disease activity who would benefit from more efficacious, but sometimes aggressive, treatments. The Rio score (Rio et al., 2009) and its modified version (Sormani et al., 2013) are scoring systems aimed at predicting treatment response in MS patients, based on the EDSS score, relapse rate, and the number of lesions. They provide a framework for stratifying patients into groups with varying likelihoods of treatment response, helping to guide treatment decisions and optimize patient outcomes. More recently, Pruenza et al. (2019) used demographic, clinical and paraclinical variables to predict individual response to fourteen different DMTs in terms of the evolution of the patient's disability measured with EDSS within 1 or 2 years, with AUC results around 0.8. The DMT type (or lack of DMT) entered the model as a variable, along with other

clinical features. Interestingly, their model consistently suggested that not treating a patient is the worst option.

2.4 Predicting conversion to definite MS or secondary progressive MS

Research into predictive modeling in MS can also take the shape of predicting the conversion from clinically isolated syndrome to definite MS, or from relapsing-remitting MS to secondary progressive MS. Zhang et al. (2019) found that features related to the shape of brain lesions on MRI (lesion volume, sphericity, surface-volume ratio) are more informative for predicting the conversion to definite MS than the well-accepted clinical criteria of lesion dissemination in space (i.e., presence of lesions in characteristic anatomic locations in the brain or spine). Pinto et al. (2020) used only clinical information to obtain predictions on conversion from the relapsing-remitting to the secondary progressive MS course, and on disease severity with rapid accumulation of disability after 6 and 10 years of progression. The most predictive features included the EDSS score and the patient's age at onset, with AUC values around 0.85. Zhao et al. (2017) employed support vector machines to build a classifier using demographic, clinical and MRI data obtained at years one and two to predict EDSS at five years follow-up; only using baseline data provided little predictive value, but clinical observations for one year improved overall sensitivity and specificity in predicting worsening cases, though values were in the 60%. A classifier incorporating short-term clinical and brain MRI data, class imbalance corrective measures, and classification costs may be a promising means to predict MS disease course, and for selection of patients suitable for more aggressive treatment regimens.

3 Breast Cancer Use Case

3.1 Synoptic pathology report examination

Pathology Report Analysis

In the context of the HeKDisco breast cancer use case, we will perform pathology report analysis by extracting the parameters with the help of natural language processing-based algorithms. Thus, the algorithm will be able to extract the pathology parameters by overcoming textual imperfections such as typos.

In the literature, Yazdani et al (2020) clearly saw the lack of automated misspelling detection and correction systems in medical and healthcare services. This work describes the development of an automated misspelling detection and correction system for radiology and ultrasound free texts in the Persian language. They used the N-gram language model and three different free text types related to abdominal and pelvic ultrasound, head and neck ultrasound, and breast ultrasound reports. Their systems achieved a detection performance of up to 90.29% in radiology and ultrasound free texts with a correction accuracy of 88.56%. In addition, Thiebaut and al. (2017) presented a text structuring technique (about breast cancer texts) that does not require to complete an existing corpus with local or new notations. They have developed an approach to reduce the semantic complexity of documents, thereby making it workable for search and information retrieval algorithms. To do so they have created a synonym dictionary specific to their corpus using the recent Word2Vec algorithm and also they use typographical errors corrections (for this they relied on the Levenshtein edit distance between pairs of words). Also, Pogrebnoy (2023) fine-tuned a version of the ruRoberta-large model on a specially collected dataset of over 30,000 medical anamneses in Russian. This model was created for correcting typos in medical histories using BERT models as a ranking of candidates.

From the technical perspective, we are planning to use transformers like BERT, clinicalBert models and/or Masked Language Models on the breast cancer pathological texts for extracting essential parameters while ignoring textual imperfections such as typos. In our algorithms, deep learning models will be trained on a large breast cancer dataset with a large corpus of words. The model will be trained from state of the art transformers like BERT series for fine tuning.

3.2 Double-check for the pathology outcomes

Mitosis Detection

Mitotic count has a key role to assess the tumor aggressiveness and proliferation in breast cancer. We conducted literature research about mitosis detection algorithms. In this context, there exists a comprehensive challenge, called MIDOG (Mitosis Domain Generalization), which was held in 2021 and 2022. The MIDOG-2021 dataset includes breast samples from 4 different scanners. The MIDOG-2022

challenge extended this dataset including samples from 6 different tumor types. Thus, the organizers of this challenge encourage the participants to develop a robust and generalizable model to different domains. F1-score is considered as the main metric to compare the mitotic figure performance. Best performing algorithm of MIDOG-2021 has an F1-score of 0.75 (Yang et al, 2021). Best performing algorithm of MIDOG-2022 has an F1-score of 0.76 (Jahanifar et al, 2021). The most common methods in the challenge are segmentation- and object detection-based deep learning algorithms. Best performing algorithm of MIDOG-2021 performs swapping the low-frequency information of images from different scanners in the Fourier domain. Thereby, the method enables to generate images with different styles since high frequency information has mitosis signals while low frequency information has background signals. Best performing algorithm of MIDOG-2022 developed a two-staged algorithm which consists of segmentation followed by classification. Additionally, they utilized ensemble learning to improve the performance.

HER2 scoring

HER2 scoring is crucial to make a decision whether the patient will take chemotherapy or not. This score might be 0, 1, 2 and 3. Kabakçı et al. (2021) proposed a traditional image processing algorithm to detect membrane and nucleus structures. Subsequently, they trained a classifier by using the histogram information of those structures. They achieved an F1-score of 91.81%. Tewary et al. (2020) developed a machine learning based classifier, called AutoIHC-Analyzer, to predict the HER2 score. Accordingly, they extract the features after detecting negative cells, complete and incomplete membrane cells. They achieved an F1-score of 93%.

Foundation Models

A new paradigm has emerged in artificial intelligence based on very large deep neural networks trained on enormous datasets, termed foundation models. Vorontsov et al. (2023) created Virchow, a 632 million parameter deep neural network foundation model for computational pathology. Using self-supervised learning, Virchow is trained on 1.5 million hematoxylin and eosin-stained whole slide images from diverse tissue groups. Virchow outperforms state-of-the-art systems both on internal datasets drawn from the same population as the pretraining data as well as external public datasets.

State of the art studies in digital pathology focus on developing visual language models (VLM) which can be applied to downstream visual recognition tasks without fine-tuning. The VLM pre-training is usually guided by certain vision-language objectives that enable it to learn image-text correspondences from the large-scale image-text pairs (Zhang et al, 2024). Lu et al. (2023) introduced CONtrastive learning from Captions for Histopathology (CONCH), a visual-language foundation model developed using diverse sources of histopathology images, biomedical text, and

notably over 1.17 million image-caption pairs via task-agnostic pretraining. Evaluated on a suite of 13 diverse benchmarks, CONCH can be transferred to a wide range of downstream tasks involving either or both histopathology images and text, achieving state-of-the-art performance on histology image classification, segmentation, captioning, text-to-image and image-to-text retrieval.

3.3 3-month mortality in metastatic disease

Month Mortality Prediction of Patients with Metastatic Tumors

Zachariah et al. (2022) proposed a machine learning model to predict 3-month mortality in patients with metastatic solid tumors. They trained a gradient-boosted trees binary classifier via the XGBoost library with observations from 28 484 deceased and alive patients and 493 features from demographic characteristics, laboratory test results, flowsheets, and diagnoses collected from the EHR between January 1, 2013, and April 24, 2019. The results of this study suggest the potential for a machine learning model trained with electronic health record data to support oncologists in prognostication and clinical decision-making to improve end-of-life care.

4 Sepsis Use Case

Sepsis is a potentially life-threatening condition that arises when the body's response to an infection causes widespread inflammation, leading to organ dysfunction. It can affect anyone with an infection, ranging from mild to severe. Sepsis progresses through stages, beginning with infection, then leading to sepsis, severe sepsis, and ultimately septic shock if left untreated. Timely recognition and treatment are critical for improving outcomes and reducing sepsis mortality rates (Ackerman et al., 2021). Machine learning (ML) models play a vital role in detecting sepsis by leveraging vast amounts of patient data to identify patterns and markers indicative of the condition. These models can analyze data from various sources, including electronic health records, vital sign monitors, laboratory results, and even wearable devices (Fleuren et al., 2020). By examining trends and anomalies in this data, machine learning algorithms can flag patients who are at risk of developing sepsis or who already have it but may not yet show obvious symptoms. Additionally, machine learning can help predict the progression of sepsis and guide treatment decisions, such as the need for antibiotics or fluid resuscitation (Kijpaisalratana et al., 2024). By integrating machine learning into clinical workflows, healthcare teams can potentially identify sepsis earlier, initiate appropriate interventions sooner, and ultimately improve patient outcomes.

4.1 Machine learning applications for sepsis

Several applications of machine learning in sepsis detection have been proposed, aiming to enhance early recognition, prediction, and management of the condition. Machine learning (ML) holds significant promise for revolutionizing sepsis detection and management across healthcare settings. ML algorithms analyze vast amounts of patient data, including vital signs, laboratory results, and clinical notes, to identify early signs of sepsis and generate real-time alerts for healthcare providers (Valik et al., 2023). These systems facilitate early intervention, potentially preventing the progression of sepsis to more severe stages. Additionally, ML models stratify patients based on their risk of developing sepsis or experiencing complications (Gholamzadeh et al., 2023), providing personalized risk scores to guide clinical decision-making. Predictive models forecast the progression of sepsis, enabling clinicians to anticipate deterioration and adjust treatment strategies accordingly. Furthermore, ML techniques categorize patients into subgroups based on clinical characteristics (Papathanakos et al., 2023), facilitating targeted interventions tailored to specific phenotypes. Integration with electronic health records (EHR) ensures seamless data analysis and decision support at the point of care. Remote monitoring and wearable devices equipped with ML algorithms enable continuous surveillance of early indicators of sepsis, enhancing early detection and intervention. Overall, ML-based approaches offer promising avenues for improving sepsis outcomes by leveraging data analytics to enhance clinical decision-making and patient care.

4.2 Machine learning techniques used in sepsis applications

In the field of sepsis prediction, there are numerous studies on selecting the correct features. A study has been examined where the selection of the most important features for the dependent variable is made using Generalized Linear Model (Rahman et al., 2023). This is one of the methods used in the HeKDisco project to evaluate features after data cleaning.

For predicting sepsis and sepsis mortality, multilayer perceptrons, neural networks with deep learning, Long Short-Term Memory (LSTM), and Recurrent Neural Networks (RNN) are frequently employed (Rafiei et al., 2021; Sharma et al., 2022). Feedforward and back-propagation learning algorithms are implemented in these models. The utilization of a fixed learning rate back-propagation algorithm in feedforward neural networks can lead to numerous complications owing to the configuration of their hidden layers and the quantity of neurons. Overfitting, insufficient training, and the inability to attain the global minimum solution are a few of these issues. Gradient-based methods utilize the gradient from a given point in the solution space to determine the global minimum; however, they are susceptible to becoming entangled in local minima. An alternative approach for generating optimal solutions by avoiding local minima is presented in the study by Kaya et al. (2023), which employs meta-heuristic algorithms.

A study at Washington University (Ewig et al., 2023) proposed a multi-subset approach to fill in missing data entered within a 6-hour timeframe by making early predictions. Furthermore, the study generated a sequence of delta and statistical values to capture the temporal trend of change, potentially of similar importance for various patients. Amrollahi et al. (2020) used pre-trained neural language representation models to improve early sepsis detection using the MIMIC-III dataset, achieving a ROC-AUC value of 0.84. Kaji et al. (2019) developed a model that predicted the onset of sepsis on the same day and the following day, achieving a ROC-AUC of 0.87. Similarly, Nemati et al. (2018) introduced an algorithm with a ROC-AUC of 0.85. Another significant study is the one by Scherpf et al. (2019), who developed a model capable of predicting the onset of sepsis three hours in advance with a ROC-AUC of nearly 0.8.

5 Conclusions

The state-of-the-art analysis for the distinct use cases included in the HeKDisco project show that there can be a positive impact of machine learning techniques in enhancing diagnostic accuracy, predicting disease progression, and optimizing treatment strategies. Advanced algorithms, such as deep learning and ensemble methods, can be leveraged to extract meaningful insights from complex medical data. Furthermore, the use of large-scale datasets and distributed infrastructure underscores the potential for future development of more personalized medical device software products and services to support effective healthcare interventions.

References

Ackerman, M. H., Ahrens, T., Kelly, J., & Pontillo, A. (2021). Sepsis. *Critical Care Nursing Clinics*, 33(4), 407-418.

Amrollahi, F., Shashikumar, S. P., Razmi, F., & Nemati, S. (2020). Contextual embeddings from clinical notes improves prediction of sepsis. In *AMIA annual symposium proceedings (Vol. 2020, p. 197)*. American Medical Informatics Association.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. 3D U-net: learning dense volumetric segmentation from sparse annotation. In *MICCAI International Conference (2016)*, Springer, pp. 424–432.

Coll, L., Pareto, D., Carbonell-Mirabent, P., Cobo-Calvo, Á., Arrambide, G., Vidal-Jordana, Á., Comabella, M., Castelló, J., Rodríguez-Acevedo, B., Zabalza, A., et al. (2023). Deciphering multiple sclerosis disability with deep learning attention maps on clinical MRI. *NeuroImage. Clinical*, 38, 103376.
<https://doi.org/10.1016/j.nicl.2023.103376>

Cruciani, F., Brusini, L., Zucchelli, M., Retuci Pinheiro, G., Setti, F., Boscolo Galazzo, I., Deriche, R., Rittner, L., Calabrese, M., & Menegaz, G. (2021). Interpretable deep learning as a means for decrypting disease signature in multiple sclerosis. *Journal of neural engineering*, 18(4), 10.1088/1741-2552/ac0f4b. <https://doi.org/10.1088/1741-2552/ac0f4b>

Danelakis, A., Theoharis, T., & Verganelakis, D. A. (2018). Survey of automated multiple sclerosis lesion segmentation techniques on magnetic resonance imaging. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 70, 83–100.
<https://doi.org/10.1016/j.compmedimag.2018.10.002>

D'hooghe, M. B., Gielen, J., Van Remoortel, A., D'haeseleer, M., Peeters, E., Cambron, M., De Keyser, J., & Nagels, G. (2019). Single MRI-Based Volumetric Assessment in Clinical Practice Is Associated With MS-Related Disability. *Journal of magnetic resonance imaging : JMRI*, 49(5), 1312–1321. <https://doi.org/10.1002/jmri.26303>

Eshaghi, A., Young, A. L., Wijeratne, P. A., Prados, F., Arnold, D. L., Narayanan, S., Guttman, C. R. G., Barkhof, F., Alexander, D. C., Thompson, A. J., Chard, D., & Ciccarelli, O. (2021). Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. *Nature communications*, 12(1), 2078.
<https://doi.org/10.1038/s41467-021-22265-2>

Ewig, K., Lin, X., Stewart, T., Stern, K., O'Keefe, G., Teredesai, A., & Hu, J. (2023). Multi-subset approach to early sepsis prediction. *arXiv preprint arXiv:2304.06384*.

Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., et al. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. *Brain: a journal of neurology*, 142(7), 1858–1875. <https://doi.org/10.1093/brain/awz144>.

Fleuren, L. M., Klausch, T. L., Zwager, C. L., Schoonmade, L. J., Guo, T., Roggeveen, L. F., ... & Elbers, P. W. (2020). Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. *Intensive care medicine*, 46, 383-400.

Gholamzadeh, M., Abtahi, H., & Safdari, R. (2023). Comparison of different machine learning algorithms to classify patients suspected of having sepsis infection in the intensive care unit. *Informatics in Medicine Unlocked*, 38, 101236.

Jahanifar, M., Shepard, A., Zamanitajeddin, N., Bashir, R. S., Bilal, M., Khurram, S. A., ... & Rajpoot, N. (2021, September). Stain-robust mitotic figure detection for the mitosis domain generalization challenge. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 48-52). Cham: Springer International Publishing.

Kabakçı, K. A., Çakır, A., Türkmen, İ., Töreyn, B. U., & Çapar, A. (2021). Automated scoring of CerbB2/HER2 receptors using histogram based analysis of immunohistochemistry breast cancer tissue images. *Biomedical Signal Processing and Control*, 69, 102924. <https://doi.org/10.1016/j.bspc.2021.102924>

Kaji, D. A., Zech, J. R., Kim, J. S., Cho, S. K., Dangayach, N. S., Costa, A. B., & Oermann, E. K. (2019). An attention based deep learning model of clinical events in the intensive care unit. *PloS one*, 14(2), e0211057.

Kaya, U., Yılmaz, A., & Aşar, S. (2023). Sepsis prediction by using a hybrid metaheuristic algorithm: A novel approach for optimizing deep neural networks. *Diagnostics*, 13(12), 2023.

Kijpaisalratana, N., Saoraya, J., Nhuboonkaew, P., Vongkulbhisana, K., & Musikatavorn, K. (2024). Real-time machine learning-assisted sepsis alert enhances the timeliness of antibiotic administration and diagnostic accuracy in emergency department patients with sepsis: a cluster-randomized trial. *Internal and Emergency Medicine*, 1-10.

Kurtzke J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*, 33(11), 1444–1452. <https://doi.org/10.1212/wnl.33.11.1444>

Lu, Ming & Chen, Bowen & Williamson, Drew & Chen, Richard & Liang, Ivy & Ding, Tong & Jaume, Guillaume & Odintsov, Igor & Zhang, Andrew & Le, Long & Gerber,

Georg & Parwani, Anil & Mahmood, Faisal. (2023). Towards a Visual-Language Foundation Model for Computational Pathology.

Marzullo, A., Kocevar, G., Stamile, C., Durand-Dubief, F., Terracina, G., Calimeri, F., & Sappey-Mariniere, D. (2019). Classification of Multiple Sclerosis Clinical Profiles via Graph Convolutional Neural Networks. *Frontiers in neuroscience*, 13, 594.
<https://doi.org/10.3389/fnins.2019.00594>

Nemati, S., Holder, A., Razmi, F., Stanley, M. D., Clifford, G. D., & Buchman, T. G. (2018). An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Critical care medicine*, 46(4), 547-553.

Papathanakos, G., Andrianopoulos, I., Xenikakis, M., Papathanasiou, A., Koulenti, D., Blot, S., & Koulouras, V. (2023). Clinical sepsis phenotypes in critically ill patients. *Microorganisms*, 11(9), 2165.

Pinto, M. F., Oliveira, H., Batista, S., Cruz, L., Pinto, M., Correia, I., Martins, P., & Teixeira, C. (2020). Prediction of disease progression and outcomes in multiple sclerosis with machine learning. *Scientific reports*, 10(1), 21038.
<https://doi.org/10.1038/s41598-020-78212-6>

Pogrebnoy D., Hugging Face website, accessed November 1, 2023.
<https://huggingface.co/DmitryPogrebnoy/MedRuRobertaLarge>

Popescu, V., Agosta, F., Hulst, H. E., Sluimer, I. C., Knol, D. L., Sormani, M. P., Enzinger, C., Ropele, S., Alonso, J., Sastre-Garriga, J., et al. (2013). Brain atrophy and lesion load predict long term disability in multiple sclerosis. *Journal of neurology, neurosurgery, and psychiatry*, 84(10), 1082–1091. <https://doi.org/10.1136/jnnp-2012-304094>

Pruenza, C., Solano, M. T., Díaz, J., Arroyo, R., & Izquierdo, G. (2019). Model for prediction of progression in multiple sclerosis. *IJIMAI*, 5(6), 47-53.
<https://doi.org/10.9781/ijimai.2019.06.005>

Rafiei, A., Rezaee, A., Hajati, F., Gheisari, S., & Golzan, M. (2021). SSP: Early prediction of sepsis using fully connected LSTM-CNN model. *Computers in biology and medicine*, 128, 104110.

Rahman, M. A., Ibrahim, A. A., & Tumian, A. (2023, February). Feature Selection using Generalized Linear Model for Machine Learning-based Sepsis Prediction. In *2023 International Conference on Advances in Intelligent Computing and Applications (AICAPS)* (pp. 1-6). IEEE.

Rakić, M., Vercruyssen, S., Van Eyndhoven, S., de la Rosa, E., Jain, S., Van Huffel, S., Maes, F., Smeets, D., & Sima, D. M. (2021). *icobrain ms 5.1: Combining unsupervised*

and supervised approaches for improving the detection of multiple sclerosis lesions. *NeuroImage. Clinical*, 31, 102707. <https://doi.org/10.1016/j.nicl.2021.102707>

Río, J., Comabella, M., & Montalban, X. (2009). Predicting responders to therapies for multiple sclerosis. *Nature reviews. Neurology*, 5(10), 553–560. <https://doi.org/10.1038/nrneurol.2009.139>

Roca, P., Attye, A., Colas, L., Tucholka, A., Rubini, P., Cackowski, S., Ding, J., Budzik, J. F., Renard, F., Doyle, S., et al. (2020). Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI. *Diagnostic and interventional imaging*, 101(12), 795–802. <https://doi.org/10.1016/j.diii.2020.05.009>

Scherpf, M., Gräßler, F., Malberg, H., & Zaunseder, S. (2019). Predicting sepsis with a recurrent neural network using the MIMIC III database. *Computers in biology and medicine*, 113, 103395.

Sharma, D. K., Lakhotia, P., Sain, P., & Brahmachari, S. (2022). Early prediction and monitoring of sepsis using sequential long short term memory model. *Expert Systems*, 39(3), e12798.

Smeets, D., Ribbens, A., Sima, D. M., Cambron, M., Horakova, D., Jain, S., Maertens, A., Van Vlierberghe, E., Terzopoulos, V., Van Binst, A. M., Vaneckova, M., Krasensky, J., Uher, T., Seidl, Z., De Keyser, J., Nagels, G., De Mey, J., Havrdova, E., & Van Hecke, W. (2016). Reliable measurements of brain atrophy in individual patients with multiple sclerosis. *Brain and behavior*, 6(9), e00518. <https://doi.org/10.1002/brb3.518>

Sormani, M. P., Río, J., Tintorè, M., Signori, A., Li, D., Cornelisse, P., Stubinski, B., Stromillo, M.I., Montalban, X., & De Stefano, N. (2013). Scoring treatment response in patients with relapsing multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)*, 19(5), 605–612. <https://doi.org/10.1177/1352458512460605>

Storelli L., Azzimonti M., Gueye M., Vizzino C., Preziosa P., Tedeschi G., de Stefano N., Pantano P., Filippi M., Rocca M.A. A Deep Learning Approach to Predicting Disease Progression in Multiple Sclerosis Using Magnetic Resonance Imaging. *Invest. Radiol.* 2022;57(7):423–432. <https://doi.org/10.1097/RLI.0000000000000854>

Tewary, S., Arun, I., Ahmed, R., Chatterjee, S., & Mukhopadhyay, S. (2021). AutoIHC-Analyzer: computer-assisted microscopy for automated membrane extraction/scoring in HER2 molecular markers. *Journal of microscopy*, 281(1), 87–96. <https://doi.org/10.1111/jmi.12955>

Thiebaut, N., Simoulin, A., Neuberger, K., Ibnouhsein, I., Bousquet, N., Reix, N., ... & Mathelin, C. (2017). An innovative solution for breast cancer textual big data analysis. *arXiv preprint arXiv:1712.02259*.

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, et al. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet. Neurology*, 17(2), 162–173. [https://doi.org/10.1016/S1474-4422\(17\)30470-2](https://doi.org/10.1016/S1474-4422(17)30470-2)

Tousignant, A., Lemaître, P., Doina, P., Arnold, D.L., Arbel, T. (2019). Prediction of Disease Progression in Multiple Sclerosis Patients using Deep Learning Analysis of MRI Data . *Proc. Mach. Learn. Res.*102:483–492.
<http://proceedings.mlr.press/v102/tousignant19a.html>

Valik, J. K., Ward, L., Tanushi, H., Johansson, A. F., Färnert, A., Mogensen, M. L., ... & Naucér, P. (2023). Predicting sepsis onset using a machine learned causal probabilistic network algorithm based on electronic health records data. *Scientific reports*, 13(1), 11760.

Vorontsov, E., Bozkurt, A., Casson, A., Shaikovski, G., Zelechowski, M., Liu, S., ... & Fuchs, T. J. (2023). Virchow: A million-slide digital pathology foundation model. *arXiv preprint arXiv:2309.07778*.

Yang, S., Luo, F., Zhang, J., & Wang, X. (2021). Sk-Unet model with Fourier domain for mitosis detection. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 86-90). Cham: Springer International Publishing.

Yazdani, A., Ghazisaeedi, M., Ahmadinejad, N., Giti, M., Amjadi, H., & Nahvijou, A. (2020). Automated Misspelling Detection and Correction in Persian Clinical Text. *Journal of digital imaging*, 33(3), 555–562. <https://doi.org/10.1007/s10278-019-00296-y>

Young, A. L., Marinescu, R. V., Oxtoby, N. P., Bocchetta, M., Yong, K., Firth, N. C., Cash, D. M., Thomas, D. L., Dick, K. M., Cardoso, J., et al. (2018). Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nature communications*, 9(1), 4273.
<https://doi.org/10.1038/s41467-018-05892-0>

Zachariah, F. J., Rossi, L. A., Roberts, L. M., & Bosserman, L. D. (2022). Prospective Comparison of Medical Oncologists and a Machine Learning Model to Predict 3-Month Mortality in Patients With Metastatic Solid Tumors. *JAMA network open*, 5(5), e2214514. <https://doi.org/10.1001/jamanetworkopen.2022.14514>

Zhang, H., Alberts, E., Pongratz, V., Mühlau, M., Zimmer, C., Wiestler, B., & Eichinger, P. (2019). Predicting conversion from clinically isolated syndrome to multiple sclerosis- An imaging-based machine learning approach. *NeuroImage. Clinical*, 21, 101593.
<https://doi.org/10.1016/j.nicl.2018.11.003>

Zhang, J., Huang, J., Jin, S., & Lu, S. (2024). Vision-language models for vision tasks: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Zhao, Y., Healy, B. C., Rotstein, D., Guttman, C. R., Bakshi, R., Weiner, H. L., Brodley, C. E., & Chitnis, T. (2017). Exploration of machine learning techniques in predicting multiple sclerosis disease course. PloS one, 12(4), e0174866.
<https://doi.org/10.1371/journal.pone.0174866>