JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Survey on the usage of Multimodel Data in ML/DL: Application and Challenges

Hetvi Shah [1], Kavita Pandya [2]
B.E Computer engineering, Samarth Group of Engineering[1]
Research scholar, Gujarat Technological University, Ahmedabad[2]
Gujarat Technological University, Himatnagar, India[1 2]

Abstract: Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without being explicitly programmed automatically and deep learning is a branch of machine learning that is based on artificial neural network architecture. An artificial neural network, or ANN, uses layers of interconnected nodes called neurons that work together to process and learn from the input data. however with the help of multimodel data in ML/DL make task more efficient and easy to manipulate with bulk data. Multimodal learning in machine learning is a type of learning where the model is trained to understand and work with multiple forms of input data. Multimodal machine learning is the study of computer algorithms that learn and improve performance through the use of multimodal datasets in various real-world applications.

Keywords: AI, ML and DL, Multimodel data applications

1. Introduction:

Machine learning [5] is an AI technique that teaches computers to learn from experience.Machine learning algorithms use computational methods to “learn” information directly from data without relying on a predetermined equation as a model. The algorithms adaptively improve their performance as the number of samples available for learning increases. Deep learning is a specialized form of machine learning.Machine learning uses two types of techniques: supervised learning, which trains a model on known input and output data so that it can predict future outputs, and unsupervised learning, which finds hidden patterns or intrinsic structures in input data. Machine learning is a growing technology that enables computers to learn automatically from past data. Machine learning uses various algorithms for building mathematical models and making predictions using historical data or information. Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommended systems, and many more.This machine learning course gives you an introduction to machine learning and the wide range of machine learning techniques such as Supervised, Unsupervised, and Reinforcement learning.Machine learning deals with the issue of how to build programs that improve their performance on some tasks through experience. Machine learning algorithms have proven to be of great practical value in a variety of application domains. They are particularly useful for poorly understood problem domains where little knowledge exists for humans to develop effective algorithms; domains where there are large databases containing valuable implicit regularities to be discovered; or domains where programs must adapt to changing conditions. Not surprisingly, the field of software engineering turns out to be a fertile ground where many software development and maintenance tasks could be formulated as learning problems and approached in terms of learning algorithms.

Comparison:

Artificial Intelligence (AI) [1]

AI is a broad field that includes ML and DL. For example, the Oxford English Dictionary defines AI as: “The theory and development of computer systems able to perform tasks normally requiring human intelligence.” However, one of my favorite definitions is by François Chollet, creator of Keras, who defined it in simplistic terms. He described AI as “the effort to automate intellectual tasks normally performed by humans”.

Machine Learning (ML) [5]-[2]

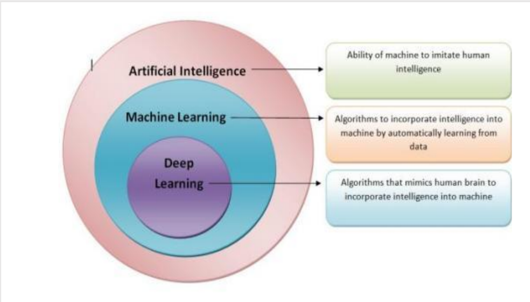
One of the pioneers of ML, Arthur Samuel, defined it as a “field of study that gives computers the ability to learn without being explicitly programmed”. ML is a subset of AI which means all ML algorithms are classified as being part of AI. However, it doesn’t work the other way, and it is important to note that not all AI based-algorithms are ML. This is analogous to how a square is a rectangle, but not every rectangle is a square.

Deep Learning(DL) [4]

DL is ML taken to the next level. It is a subset of ML that is inspired by how human brains work. Typically, when people use the term deep learning, they refer to deep artificial neural networks. DL effectively teaches computers to do what humans naturally do.

Multi model data [3]-[6]

Multimodal data refers to data that combines multiple modes or types of information. These modes can include text, images, audio, video, and other forms of data. Multimodal data analysis involves extracting meaningful insights by integrating and analyzing these different types of data.Multimodal data has broad applications across various fields,education,including autonomous-vehicles,healthcare Communication,NighttimeDetection,ClinicalPsychology,Nonsevere COVID-19,robotics, and natural language processing. By combining different types of data, researchers and practitioners can unlock new insights and improve decision-making in complex systems.



(1.1 Comparison of ML,DL,AI)

2. Survey of application using multi-model data

1. Multimodal Data-Driven Smart Healthcare Systems

The multimodal data-driven approach has emerged as an important driving force for smart healthcare systems with applications ranging from disease analysis to triage, diagnosis and treatment. A smart healthcare system [9] necessitates new demands for data management and decision-making, which development of medical services using artificial intelligence and new transformations in the healthcare industry. In a survey of existing techniques which includes not only state-of-the-art methods but also the most recent trends in the field. In particular, this review focuses on the types of decision-making processes used in smart healthcare systems. Firstly, approaches that utilize multimodal association mining with fine-grained data semantics in smart healthcare systems are introduced.Second approaches for multimodal data fusion and cross-border association that have been employed in developing smart healthcare systems. Focus specifically on the use of the panoramic decision framework, interactive decision making, and intelligent decision support systems.The multimodal data-driven smart healthcare system is becoming an important technical carrier and driving force for the entire process of healthcare, including consultation, triage, diagnosis and treatment.The clinical performance of smart healthcare systems must be continually improved by further developing these theoretical algorithms such as graph neural network,deep learning-based intelligent decision-making engines, and multimodal methods.

Datasets:

College of some medical imaging applications in which deep learning has achieved state-of-the-art results.skin cancer classification, macular degeneration and diabetic retinopathy classification, prostate cancer detection , MRI brain feature extraction, mitosis detection in breast cancer, indentification of fetal standard scan planes, polyp detection, evaluating capsule endoscopy videos, and thoracolumbar fractures detection. In healthcare, big data can be collected from the following four sources: clinical information systems, personal mobile or wearable healthcare devices, Internet Of Things (IoTs) and open medical data.

3. Multimodal data integration using machine learning improves the risk of high-grade serous ovarian cancer

Patients with high-grade serous ovarian cancer suffer poor prognosis and variable response to treatment Known prognostic factors for this disease include homologous recombination deficiency status, age, pathological stage and residual disease status after debulking surgery.Radiomic feature log HRs, radiomic model forest plot and radiomic score versus OS and c-Indices survery has highlighted important prognostic information captured in computed tomography and histopathological specimens, which can be exploited through machine learning. However, little is known about the capacity of combining features from these disparate sources to improve prediction of treatment response. Here,they assembled a multi modal dataset of 444 patients with primarily late-stage high-grade serous ovarian cancer [8] and discovered quantitative features, such as tumor nuclear size on staining with hematoxylin and eosin and omental texture on contrast-enhanced computed tomography, associated with prognosis.Survey found that these features contributed complementary prognostic information relative to one another and clinicogenomic features. By fusing histopathological, radiologic and clinicogenomic machine-learning models, They demonstrate a promising path toward improved risk stratification of patients with cancer through multimodal data integration.To collect H&E imaging, they reviewed the electronic health record to find associated pathology cases with peritoneal lesions (primarily omental) and expert pathologists reviewed the slides to select high-quality specimens for digitization.They also reviewed the institutional data repository for scanned slides associated with the diagnostic biopsy and included those containing tumors. All H&E imaging was carried out before treatment.They subsequently reviewed the associated CE-CT scans for the following the inclusion criteria: (1) intravenous contrast-enhanced images acquired in the portal venous phase, (2) absence of streak artifacts or motion-related image blur obscuring lesion(s) of interest and (3) adequate signal to noise ratio (Supplementary Table 7). All CE-CT imaging was carried out before treatment. All CT scans were available in the digital imaging and communications in medicine (DICOM) format through our institutional picture archiving and communication system.Corroborating this, absolute Kendall rank correlation coefficient values were low between individual modalities (<0.14), demonstrating that the radiomic and histopathological models ordered patients differently as compared to the genomic model and to one another. The same two risk groups identified by the model in the test set also showed significantly different progression-free survival (PFS) ($P = 0.040$). Finally, as an orthogonal validation, the inferred risk of all models except the genomic and genomic–histopathological models associated with pathological chemotherapy response score (CRS) in the training set, including the GRH model. The test set had only 21 patients with known CRS and only HRD status exhibited statistically significantly different distributions of CRS by the Mann–Whitney U-test in the test set.

Datasets:

Radiomic feature log HRs, radiomic model forest plot and radiomic score versus OS and c-Indices.Multimodal c-Indices, forest plots, multimodal risk score versus OS/PFS, unimodal score comparison and CRS versus multimodal score and Kendall's tau correlation.CRS categories versus individual model scores in the test set.Uncorrected histopathological feature log HRs.

4. A Multimodality Machine Learning to Differentiate Severe and Nonsevere COVID-19 Model Development and Validation

Using [11] clinical and laboratory results independently as input, the random forest models achieved >90% and >95% predictive accuracy, respectively. The importance scores of the input features were further evaluated, and the top 5 features from each modality were identified (age, hypertension, cardiovascular disease, gender, and diabetes for the clinical features modality, and dimerized plasmin fragment D, high sensitivity troponin I, absolute neutrophil count, interleukin 6, and lactate dehydrogenase for the laboratory testing modality, in descending order). Using these top 10 multimodal features as the only input instead of all 52 features combined, the random forest model was able to achieve 97% predictive accuracy.

5. Multi-Modal Data and Machine Learning to Improve Cardiovascular Disease Care

These data were utilized in silos, new machine learning (ML) and deep learning (DL) technologies enable the integration of these data sources to produce multi-modal insights. Data fusion, which integrates data from multiple modalities using ML and DL techniques, has been of growing interest in its application to medicine. In this paper, Review the state-of-the-art research that focuses on how the latest techniques in data fusion are providing scientific and clinical insights specific to the field of cardiovascular medicine. With these new data fusion capabilities, clinicians and researchers alike will advance the diagnosis and treatment of cardiovascular diseases (CVD) [7] to deliver more timely, accurate, and precise patient care.Multimodal data fusion and machine learning in cardiovascular medicine is an exciting field of research, though, there are still very few use cases to date. Using data from multiple modalities offers the promise of improved AI technology whereby the weaknesses of each type of health care data can be addressed through different data combinations. However, algorithms used to analyze multiple data modalities may be too complex, too difficult to implement, and too slow to fit into a time frame that makes them usable in a clinical work environment. Furthermore, a focus on data quality will be essential to prevent exponentially propagating errors when combining data. Future research should focus on streamlined methods for data integration, best practices for evaluating model gain from different types of data, and prospective study designs to validate clinical utility.Multimodal data refers to data that spans different types and contexts. Methods used to fuse multimodal data fundamentally aim to integrate the data with values of different scales and distributions into a global feature space in which data can be represented in a more uniform manner. This uniformity can then be leveraged for tasks such as prediction and classification. For example, data from large biobanks such as the UK biobank, the Million Veterans Program, and the National Institutes of Health All of Us initiative contain patient-specific genomic data, imaging studies, and phenotypic data from EMR and questionnaires. Each of these data types can be fused to predict cardiovascular disease prognosis, improve identification of unique subgroups, and predict response to treatment. The hope is that more accurate models can be built with multiple types of data than if only one type of data were utilized.Some patients may require a greater focus on social determinants of health to improve outcomes in addition, to adequate medical management. With this in mind, Flores et al. aimed to evaluate whether multimodal data could help provide insights into different cardiovascular phenotypes that might lend themselves to different clinical approaches. Previous work in the domain of cardiovascular phenotyping was described by Shah et al., who demonstrated that unsupervised learning techniques such as hierarchical clustering can be used to identify clinically meaningful subgroups of patients with CHF. While this helped establish unsupervised learning as a useful way to identify clinical subgroups that may benefit from different therapies, Shah et al. were limited by the data they could use. With traditional clustering models, data are typically required to be in the same format (numerical, ordinal, or categorical). Instead, Flores et al. aimed to combine genetic, imaging, demographic, clinical, and lifestyle data to identify cardiovascular disease subgroups using unsupervised methods.

Dataset:

Electronic medical records (EMR), radiology images, and genetic repositories.

6. Multimodal Fusion of Brain Imaging Data To Finding the Missing Link in Complex Mental Illness

It is becoming increasingly clear that combining multimodal brain imaging data [10] provides more information for individual subjects by exploiting the rich multimodal information that exists. However, the number of studies that do true multimodal fusion is still remarkably small given the known benefits. In part, this is because multimodal studies require broader expertise in collecting, analyzing, and interpreting the results than do unimodal studies.Survey start by introducing the basic reasons why multimodal data fusion is important and what it can do and, importantly, how it can help us avoid wrong conclusions and help compensate for imperfect brain imaging studies. Challenges that need to be confronted for such approaches to be more widely applied by the community. Review of the diverse studies that have used multimodal data fusion as well as provide an introduction to some of the existing analytic approaches. Finally,By the discuss some up-and-coming approaches to multimodal fusion including deep learning and multimodal classification that show considerable promise. Multimodal data fusion is rapidly growing, but it is still underutilized. The complexity of the human brain coupled with the incomplete measurement provided by existing imaging technology makes multimodal fusion essential to mitigate misdirection and hopefully provide a key to finding the missing links in complex mental illness."Multimodal" is a widely used phrase in the context of brain imaging studies. Collecting multiple modalities of magnetic resonance imaging (MRI) data from the same individual has been popular in brain imaging studies. There is increasing evidence that multimodal brain imaging studies can help provide a more complete understanding of the brain and its disorders, for example it can inform us about how brain structure shapes brain function, in which way they are impacted by psychopathology and which functional or structural aspects of physiology could drive human behavior and cognition.There are multiple studies demonstrating the combination of structural and functional data can improve brain disease classification. A strong demonstration of this is found in a recent multimodal classification challenge for schizophrenia versus controls using GM and rest fMRI connectivity received over 2000 submissions, most of which were able to achieve greater than 80% accuracy. Dai et al.proposed an automatic classification framework which integrated multimodal image features using multi-kernel learning (MKL) for predicting attention deficit/hyperactivity disorder versus controls finding a very high classification performance. Using an ensemble feature selection strategy and an advanced support vector machine approach, Sui et al.Resting-state fMRI, EEG and sMRI data to classify schizophrenia from healthy controls and achieved the best performance with 91% accuracy compared to using a single modality. By adopting Gaussian process classifiers to evaluate the prognostic value of neuroimaging data and clinical characteristics, Schmaal et al. discovered that prediction of the naturalistic course of depression over 2 years is improved by considering different task contrasts or data sources, especially those derived from neural responses to emotional facial expressions. Finally, Pettersson-Yeo et al. used a multimodal SVM approach to examine the ability of sMRI, fMRI, dMRI and cognitive data to differentiate between ultra-high-risk (UHR) and first-episode (FEP) psychosis at the single-subject level, supporting clinical development of SVM to help inform identification of FEP and UHR. These findings strongly suggest that multi-modal classification facilitated by advanced modeling techniques can provide more accurate and early detection of brain abnormalities beyond approaches that use only a single modality.: Major advances in performance have been obtained in multiple domains, including brain imaging, via deep (multilayered) learning algorithms to capture nonlinear/higher order relationships. Recent work has shown the potential for such models in neuroimaging data and provide a framework to extend promising approaches such as linear ICA. One potential issue is that training of deep models requires extensive amounts of data. However this issue can in part be overcome by training the models with realistic simulation data. Deep models have recently made significant advances, outperforming shallow models in multiple problem domains such as image classification. Our work shows that class separation improves with deep belief network (DBN) depth while DBNs uncover hidden relations within data and thus facilitate discovery. Specifically,they investigated if classification rates improve with depth by sequentially investigating DBNs of 3 depths. Figure 8a displays 2D maps of the raw data, as well as the depth 1, 2, and 3 activations: the deeper networks place schizophrenia patients and healthy control groups further apart for both training and validation data. Another benefit of deep learning models is their ability to automatically discover high level representation, which is especially important for multimodal analysis incorporating different data types that are unlikely to have a simple linear correspondence. An example of a multimodal deep learning architecture is shown in Figure 8b. There are many interesting emerging models, for example, motivated by the concept of brain function and structure representing static images (sMRI) annotated by sequential captions of the brain (fMRI), one can build a model to translate the relationship between brain structure and brain function using a recurrent neural network.It is becoming increasingly clear that combining multimodal brain imaging data provides more information for individual subjects by exploiting the rich multimodal information that exists. However, the number of studies that do true multimodal fusion (i.e., capitalizing on joint information among modalities) is still remarkably small given the known benefits. In part, this is because multimodal studies require broader expertise in collecting, analyzing, and interpreting the results than do unimodal studies. We start by introducing the basic reasons why multimodal data fusion is important and what it can do and, importantly, how it can help us avoid wrong conclusions and help compensate for imperfect brain imaging studies.The challenges that need to be confronted for such approaches to be more widely applied by the community.They provide a review of the diverse studies that have used multimodal data fusion (primarily focused on psychosis) as well as provide an introduction to some of the existing analytic approaches. Finally,discuss some up-and-coming approaches to multimodal fusion including deep learning and multimodal classification that show considerable promise.Conclusion is that multimodal data fusion is rapidly growing, but it is still underutilized. The complexity of the human brain coupled with the incomplete measurement provided by existing imaging technology makes multimodal fusion essential to mitigate misdirection and hopefully provide a key to finding the missing link(s) in complex mental illness.

7. Deep Learning for Image Super-resolution: A Survey

Image Super-Resolution [20] is an important class of image processing techniques to enhance the resolution of images and videos in computer vision. Recent years have witnessed remarkable progress of image super-resolution using deep learning techniques. This article aims to provide a comprehensive survey on recent advances of image super-resolution using deep learning approaches. In general, they can roughly group the existing studies of SR techniques into three major categories: supervised SR, unsupervised SR, and domain-specific SR. In addition,They cover some other important issues, such as publicly available benchmark datasets and performance evaluation metrics. Finally, conclude this survey by highlighting several future directions and open issues which should be further addressed by the community in the future.Interpolation-based Upsampling Image interpolation, a.k.a. image scaling, refers to resizing digital images and is widely used by image-related applications. The traditional interpolation methods include nearestneighbor interpolation, bilinear and bicubic interpolation, Sinc and Lanczos resampling, etc. Since these methods are interpretable and easy to implement, some of them are still widely used in CNN-based SR models. Nearest-neighbor Interpolation. The nearest-neighbor interpolation is a simple and intuitive algorithm. It selects the value of the nearest pixel for each position to be interpolated regardless of any other pixels. Thus this method is very fast but usually produces blocky results of low quality. Bilinear Interpolation. The bilinear interpolation (BLI) first performs linear interpolation on one axis of the image. Since it results in a quadratic interpolation with a receptive field sized 2 × 2, it shows much better performance than nearestneighbor interpolation while keeping relatively fast speed.Bicubic Interpolation. Similarly, the bicubic interpolation (BCI) performs cubic interpolation. Compared to BLI, the BCI takes 4 × 4 pixels into account, and results in smoother results with fewer artifacts but much lower speed. In fact, the BCI with anti-aliasing is the mainstream method for building SR datasets and is also widely used in pre-upsampling SR framework. As a matter of fact, the interpolation-based upsampling methods improve the image resolution only based on its own image signals, without bringing any more information.

TABLE 1
List of public image datasets for super-resolution benchmarks.

Dataset	Amount	Avg. Resolution	Avg. Pixels	Format	Category Keywords
BSDS300 [40]	300	(435, 367)	154, 401	JPG	animal, building, food, landscape, people, plant, etc.
BSDS500 [41]	500	(432, 370)	154, 401	JPG	animal, building, food, landscape, people, plant, etc.
DIV2K [42]	1000	(1972, 1437)	2, 793, 250	PNG	environment, flora, fauna, handmade object, people, scenery, etc.
General-100 [43]	100	(435, 381)	181, 108	BMP	animal, daily necessity, food, people, plant, texture, etc.
L20 [44]	20	(3843, 2870)	11, 577, 492	PNG	animal, building, landscape, people, plant, etc.
Manga109 [45]	109	(826, 1169)	966, 011	PNG	manga volume
OutdoorScene [46]	10624	(553, 440)	249, 593	PNG	animal, building, grass, mountain, plant, sky, water
PIRM [47]	200	(617, 482)	292, 021	PNG	environments, flora, natural scenery, objects, people, etc.
Set5 [48]	5	(313, 336)	113, 491	PNG	baby, bird, butterfly, head, woman
Set14 [49]	14	(492, 446)	230, 203	PNG	humans, animals, insects, flowers, vegetables, comic, slides, etc.
T91 [21]	91	(264, 204)	58, 853	PNG	car, flower, fruit, human face, etc.
Urban100 [50]	100	(984, 797)	774, 314	PNG	architecture, city, structure, urban, etc.

Today there are a variety of datasets available for image super-resolution, which greatly differ in image amounts, quality, resolution, and diversity, etc. Some of them provide LR-HR image pairs, while others only provide HR images, in which case the LR images are typically obtained by imresize function with default settings in MATLAB (i.e., bicubic interpolation with anti-aliasing). In Table 1 we list a number of image datasets commonly used by the SR community,and specifically indicate their amounts of HR images, average resolution, average numbers of pixels, image formats, and category keywords. Besides these datasets, some datasets widely used for other vision tasks are also employed for SR, such as ImageNet, MS-COCO, VOC2012, CelebA. In addition, combining multiple datasets for training is also popular, such as combining T91 and BSDS300, combining DIV2K and Flickr2K.

Survey on recent advances in image super-resolution with deep learning.The improvement of supervised and unsupervised SR, and also introduced some domain-specific applications. Despite great success, there are still many unsolved problems.

8. Machine learning in healthcare communication

Machine learning (ML) is a study of computer algorithms for automation through experience. ML is a subset of artificial intelligence (AI) that develops computer systems, which are able to perform tasks generally having need of human intelligence. While healthcare communication is important in order to tactfully translate and disseminate information to support and educate patients and public, ML is proven applicable in healthcare [17] with the ability for complex dialogue management and conversational flexibility. In this topical review, how the application of ML/AI in healthcare communication is able to benefit humans. This includes chatbots for the COVID-19 health education, cancer therapy, and medical imaging.Machine learning is a robust and powerful digital tool that can benefit healthcare communication with a better patient care/education, faster decision making, and reduction of resources. To date, different fields in machine learning such as NLP and DNN have been studied and applied in various components of healthcare. The innovative AI-based chatbot takes an important role as a human like conversational agent between the user and service provider. This chatbot and response system has a significant impact on our healthcare system. With the continuous advancement in technology, it is expected to have more influence in the future. AI can reduce healthcare costs and make research tasks more efficient by introducing the latest advanced algorithms and is expected to assist clinical in many areas.

9. Multimodal deep learning models for early detection of Alzheimer's disease stage

Most current Alzheimer's disease (AD) and mild cognitive disorders (MCI) studies use single a data modality to make predictions such as AD stages. The fusion of multiple data modalities can provide a holistic view of AD staging analysis. Thus, we use deep learning [18] to integrally analyze imaging (magnetic resonance imaging (MRI)), genetic (single nucleotide polymorphisms (SNPs)), and clinical test data to classify patients into AD, MCI, and controls (CN). We use stacked denoising auto-encoders to extract features from clinical and genetic data, and use 3D-convolutional neural networks (CNNs) for imaging data.They developed a novel data interpretation method to identify top-performing features learned by the deep-models with clustering and perturbation analysis. Using Alzheimer's disease neuroimaging initiative (ADNI) Dataset, we demonstrate that deep models outperform shallow models, including support vector machines, decision trees, random forests, and k-nearest neighbors. In addition, we demonstrate that integrating multi-modality data outperforms single-modality models in terms of accuracy, precision, recall, and meanF1 scores.The models have identifed hippocampus, amygdala brain areas, and the Rey Auditory Verbal Learning Test (RAVLT) as top distinguished features, which are consistent with the known AD literature.In this study, use DL models to perform multimodal data fusion(i.e. imaging, EHR and genomic SNP data) for classifying patients into CN, MCI, and AD groups. We use stacked de-noising auto-encoders for EHR and SNP, and 3D convolutional neural networks (CNNs) for MRI imaging data. After the networks are separately trained for each data modality, we apply decision trees, random forests, support vectors machines, and k-nearest neighbors to conduct integrated classification on AD staging.Diagnosing patients with AD is challenging, and the prediction accuracy remains low for staging assessment. In the report on the potential of DL for multi-modal data fusion, including:Deep-models outperform shallow models for single-modality Alzheimer's stage prediction.Novel DL framework for multi-modality data fusion outperforms single-modality DL.Novel perturbation and clustering-based feature extraction assisting DL model interpretations are capable of AD stage prediction.Application of 3D convolutional neural network architecture for MRI image data benefits the AD analysis.Despite the improved performance, our study suffers from short-comings such as limited Dataset sizes.

Datasets:
Genetic data,Clinical features,imaging MRI data

10. Vehicle Detection In Traffic Surveillance Images From Daytime To Nighttime by deep learning

Vehicle detection in traffic surveillance images is an important approach to obtaining vehicle data and rich traffic flow parameters. Recently, deep learning [19] based methods have been widely used in vehicle detection with high accuracy and efficiency. However, deep learning based methods require a large number of manually labeled ground truths (bounding box of each vehicle in each image) to train the Conventional Neural Networks (CNN). In the modern urban surveillance cameras, there are already many manually labeled ground-truths in daytime images for training.
CNN, while there are little or much less manually labeled ground truths in nighttime images.Maximum use of labeled daytime images (Source Domain) to help the vehicle detection in unlabeled nighttime images (Target Domain). For this purpose, A new method based on Faster R-CNN with Domain Adaptation (DA) to improve vehicle detection at nighttime. With the assistance of DA, the domain distribution the discrepancy between Source and Target Domains is reduced.They collected a new Datasets of 2,200 traffic images (1,200 for daytime and 1,000 for nighttime) of 57,059 vehicles for training and testing CNN. In the experiment, only using the manually labeled ground truths of daytime data, Faster R- CNN obtained 82.84% as F-measure on the nighttime vehicle detection, while the proposed method (Faster R-CNN+DA) achieved 86.39% as F-measure on the nighttime vehicle detection.They propose a Faster R-CNN with Domain Adaptation method for vehicle detection 2 in the nighttime to use labeled daytime data (source domain) to help the unlabeled nighttime data 3 (target domain). For experiments,Datasets CAU-UTRGV Benchmark containing 4 2,200 labeled traffic images to test the proposed method and other comparison methods. Using 5 style transfer based domain adaptation, the deep learning based method Faster R-CNN can be 6 improved for vehicle detection in the nighttime traffic surveillance. Without using any extra 7 hardware and labor costs, the proposed method can improve the current vehicle detection in the 8 nighttime traffic surveillance.For a better vehicle detection in traffic surveillance images during nighttime,Propose to using 38 style transfer as the DA method to mitigate the domain difference between the source domain and 39 the target domain, and then train a Faster R-CNN model for nighttime vehicle detection.

Datasets : Daily Normal traffic

11. A Machine Learning for diagnosing Chronic Kidney Disease

Chronic kidney disease (CKD) is a global health problem with high morbidity and mortality rate, and it induces other diseases. Since there are no obvious symptoms during the early stages of CKD, patients often fail to notice the disease. Early detection of CKD enables patients to receive timely treatment to ameliorate the progression of this disease. Machine learning [16] models can effectively aid clinicians achieve this goal due to their fast and accurate recognition performance.Machine learning methodology for diagnosing CKD. The CKD data set was obtained from the University of California Irvine (UCI) machine learning repository, which has a large number of missing values. KNN imputation was used to fulfill in the missing values, which selects several complete samples with the most similar measurements to process the missing data for each incomplete sample. Missing values are usually seen in real-life medical situations because patients may miss some measurements for various reasons. After effectively fulfilling out the incomplete data set, six machine learning algorithms (logistic regression, random forest, support vector machine, k-nearest neighbor, naive Bayes classifier and feed forward neural network) were used to establish models. Among these machine learning models, random forest achieved the best performance with 99.75% diagnosis accuracy. By analyzing the misjudgments generated by the established models, They proposed an integrated model that combines logistic regression and random forest by using perception, which could achieve an average accuracy of 99.83% after ten times of simulation. Hence, speculated that this methodology could be applicable to more complicated clinical data for disease diagnosis.Used KNN imputation to fill in the missing values in the data set, which could be applied to the data set with the diagnostic categories are unknown.Logistic regression (LOG), RF, SVM, KNN, naive Bayes classifier (NB) and feed forward neural network (FNN) were used to establish CKD diagnostic models on the complete CKD data sets. The models with better performance were extracted for misjudgment analysis.An integrated model that combines LOG and RF by using perceptron was established and it improved the performance of the component models in CKD diagnosis after the missing values were filled by KNN imputation.

Datasets: Data set from the University of California Irvine

12. Machine Learning-Based Damage Identification Method

It is necessary to assess damage properly for the safe use of a structure and for the development of an appropriate maintenance strategy. Although many efforts have been made to measure the vibration of a structure[12] to determine the degree of damage, the accuracy of evaluation is not high enough, so it is difficult to say that a damage evaluation based on vibrations in a structure has not been put to practical use.A method to evaluate damage by measuring the acceleration of a structure at multiple points and interpreting the results with a Random Forest, which is a kind of supervised machine learning. The proposed method uses the maximum response acceleration, standard deviation, logarithmic decay rate, and natural frequency to improve the accuracy of damage assessment.A three-step Random Forest method to evaluate various damage types based on the results of these many measurements.Then, the accuracy of the proposed method is verified based on the results of a cross-validation and a vibration test of an actual damaged specimen.They used one of the machine learning methods, Random Forest, which is known for its high accuracy and generalization performance of both classification and regression.the Random Forest is divided into three steps in order to improve the accuracy. First, the degree of anomaly of the focused area was examined in the first step of the Random Forest, and then the presence of cracks was examined in the second step. Finally, the third step of Random Forest evaluated the degree of damage using the results obtained from the first and second steps.In this study, a supervised machine learning system was developed to determine the presence of cracks and the degree of damage by using features calculated from vibration test results.The performance of the developed method is investigated by LOOCV and by considering the results when the method is applied to actual damaged specimens; it is shown that the damage could be evaluated with a high accuracy. The results suggest that not only damage detection of beams, but also damage detection of various structures can be achieved by a statistical approach, which will be meaningful in the future when various types of measurement data, in large quantities, will be accumulated.

Datasets:

Image of Bridge,Image of Aluminum and Image of Damage Alloy etc.

13. Machine learning approaches for clinical psychology and psychaitry

Machine learning[13] focus on learning statistical functions from multidimensional data sets to make generalizable predictions about individuals. The goal of this review is to provide an accessible understanding of why this approach is important for future practice given its potential to augment decisions associated with the diagnosis, prognosis, and treatment of people suffering from mental illness using clinical and biological data. To this end, the limitations of current statistical paradigms in mental health research are critiqued, and an introduction is provided to critical machine learning methods used in clinical studies. A selective literature review is then presented aiming to reinforce the usefulness of machine learning methods and provide evidence of their potential. In the context of promising initial results, the current limitations of machine learning approaches are addressed, and considerations for future clinical translation are outlined.The topic of machine learning encompasses an approach to problems as much as a set of specific methods. This approach fundamentally aims to learn information from multivariate data to fulfill the pragmatic goal of research translation by predicting outcomes for individuals rather than groups. The methods are beginning to be more widely used in clinical psychology and psychiatry and offer a promising future direction for translational research and, ultimately, clinical care.A central component of the machine learning approach is the algorithm that is used to perform classification, regression, clustering, or normative modeling. Broadly, these can be separated into supervised techniques where cases are labeled unsupervised techniques where the aim is to divide an unlabeled sample into groups of related cases, and semi-supervised techniques containing labeled and unlabeled cases. The algorithms available have often been developed within the machine learning field, but as is the case for preprocessing algorithms, other popular methods are also commonly used across statistical cultures.

Datasets: Health care system,clinician

14. MACHINE LEARNING IN EDUCATION

Machine Learning (ML) is one of the most promising application areas in the field of Information Technology.Machine learning in the education area is currently very interesting to researchers and scientists [15], and it is the main focus of our study.Evaluate the possibilities of applying and using machine learning in the education area. identifies and analyses suitable literature, research papers and articles in order to determine their categorization in the field of education, to determine the current trends of using machine learning in education.We used the Systematic Literature Review approach (SLR)to collect primary studies regarding this research scope. The main aim was to classify studies and the selected studies were then analysed and categorized using the content analysis method.Machine learning is programming computers to optimize a performance criterion using example data or experience. Implementing a machine learning algorithm means implementing a model that outputs correct information given that we have provided input data.The process through which a model learns how to make sense of input data is called "model training". Training is a key concept in machine learning.

The aim of this study was to evaluate the current state of the art in the application of machine learning in the education area. Reviewing studies under the categories showed how machine-learning algorithms can help schools or faculties to reach out to students and get them the help they need to be successful as early as possible. Student retention is an essential part of many enrolment management systems. It affects university rankings, school reputation, and financial well being. Student retention has become one of the most important priorities for decision makers in higher education institutions, so there are a lot of studies in that category.By"learning" about each student, the technology can identify weaknesses and suggests ways to improve.Assessment Base on provides constant feedback to teachers, students and parents about how the student learns, the support they need and the progress they are making towards their learning goals.

15. Machine learning for the educational sciences

Machine learning (ML) provides a powerful framework for the analysis of high-dimensional Datasets by modelling complex relationships, often encountered in modern data with many variables, cases and potentially non-linear effects. The impact of ML methods on research and practical applications in the educational sciences [14] is still limited, but continuously grows, as larger and more complex datasets become available through massive open online courses (MOOCs) and large-scale investigations. The educational sciences are at a crucial pivot point, because of the anticipated impact ML methods hold on the field. To provide educational researchers with an elaborate introduction to the topic, They provide an instructional summary of the opportunities and chal-lenges of ML for the educational sciences, show how a look at related disciplines can help learning from their experiences, and argue for a philosophical shift in model evaluation. They demonstrate how the overall quality of data analysis in educational research can benefit from these methods and ML can play a decisive role in the validation of empirical models. Specifically, overview of the types of data suitable for ML gives practical advice for the application of ML methods. In each section, analytical examples and reproducible R code. Also, extensive Appendix on ML-based applications for edu-cation.ML approaches are able to deal with large amounts of variables without encountering convergence problems and to learn high-dimensional complex, non-linear relationships—this holds for digital Big Data and mobile sensor information from devices, such as smartphones, as well as large-scale studies.The ability to handle many variables and to model complex relationships, the application of ML techniques in the educational sciences can help to improve the science as a whole.

3.Conclusion

In conclusion, this survey explored how different types of data can be combined in machine learning and deep learning. It showed that using diverse data sources can be very useful in improving models and decision-making. The survey highlighted many examples of this in different fields, demonstrating the benefits of bringing together different types of information. However, it's important to note that there are still challenges to overcome in making this approach practical and effective. Overall, the survey emphasized the potential of using multi-model data to make better models and decisions.

REFERENCES

[1] M. Ergen, “What is Artificial Intelligence? Technical Considerations and Future Perception,” Anatol J Cardiol, 2019.

[2] phData,“What do AI, ML, and DL.”

[3] P. Koupil, M. Svoboda, and I. Holubova, “MM-cat: A Tool for Modeling and Transformation of Multi-Model Data using Category Theory,” in 2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C), Fukuoka, Japan: IEEE, Oct. 2021.

[4] Ongsulee, “Artificial intelligence, machine learning and deep learning,” in 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), Bangkok: IEEE, Nov. 2017.

[5] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, Eds., Machine Learning. Berlin, Heidelberg: Springer Berlin Heidelberg, 1983.

[6] J. Lu and I. Holubová, “Multi-model Databases: A New Journey to Handle the Variety of Data,” ACM Comput. Surv., vol. 52, no. 3, pp. 1–38, May 2020.

[7] S. Amal, L. Safarnejad, J. A. Omiye, I. Ghanzouri, J. H. Cabot, and E. G. Ross, “Use of Multi-Modal Data and Machine Learning to Improve Cardiovascular Disease Care,” Front. Cardiovasc. Med., vol. 9, p. 840262, Apr. 2022.

[8] K. M. Boehm et al., “Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer,” Nat Cancer, vol. 3, no. 6, pp. 723–733, Jun. 2022.

[9] Q. Cai, H. Wang, Z. Li, and X. Liu, “A Survey on Multimodal Data-Driven Smart Healthcare Systems: Approaches and Applications,” IEEE Access, vol. 7, pp. 133583–133599, 2019.

[10] V. D. Calhoun and J. Sui, “Multimodal Fusion of Brain Imaging Data: A Key to Finding the Missing Link(s) in Complex Mental Illness,” Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, vol. 1, no. 3, pp. 230–244, May 2016.

[11] Y. Chen et al., “A Multimodality Machine Learning Approach to Differentiate Severe and Nonsevere COVID-19: Model Development and Validation,” J Med Internet Res, vol. 23, no. 4, p. e23948, Apr. 2021.

[12] P. Chun, T. Yamane, S. Izumi, and N. Kuramoto, “Development of a Machine Learning-Based Damage Identification Method Using Multi-Point Simultaneous Acceleration Measurement Results,” Sensors, vol. 20, no. 10, p. 2780, May 2020.

[13] D. B. Dwyer, P. Falkai, and N. Koutsouleris, “Machine Learning Approaches for Clinical Psychology and Psychiatry,” Annu. Rev. Clin. Psychol., vol. 14, no. 1, pp. 91–118, May 2018.

[14] S. Hilbert et al., “Machine learning for the educational sciences,” Review of Education, vol. 9, no. 3, Oct. 2021.

[15] D. Kucak, V. Juricic, and G. Dambic, “Machine Learning in Education - a Survey of Current Research Trends,” in DAAAM Proceedings, B. Katalinic, Ed., 1st ed.DAAAM International Vienna, 2018.

[16] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, “A Machine Learning Methodology for Diagnosing Chronic Kidney Disease,” IEEE Access, vol. 8, pp. 20991–21002, 2020.

[17] S. Siddique and J. C. L. Chow, “Machine Learning in Healthcare Communication,” Encyclopedia, vol. 1, no. 1, pp. 220–239, Feb. 2021.

[18] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, “Multimodal deep learning models for early detection of Alzheimer’s disease stage,” Sci Rep, vol. 11, no. 1, p. 3254, Feb. 2021.

[19] Jinlong Ji, Zhigang Xu, Hongkai Yu, Lan Fu, and Xuesong Zhou, “Domain Adaptation For Vehicle Detection In Traffic Surveillance Domain Adaptation For Vehicle Detection In Traffic Surveillance Images From Daytime To Nighttime Images From Daytime To Nighttime,” Jul. 2019.

[20] Z. Wang, J. Chen, and S. C. H. Hoi, “Deep Learning for Image Super-resolution: A Survey,” 2019.

[21] Cong Feng,Mingjian Cui,Bri-Mathias Hodge,Jie Zhang,”A Data-Driven Multi-Model Methodology with Deep Feature Selection for Short-Term Wind Forecasting”,2017.

[22] Liang Xue,Dongxiao Zhang “A multimodel data assimilation framework via the ensembleKalman filter”,2014.