Automated decision-making (ADM) is the use of data, machines and algorithms to make decisions in a range of contexts, including public administration , business, health, education, law, employment, transport, media and entertainment, with varying degrees of human oversight or intervention. ADM may involve large-scale data from a range of sources, such as databases, text, social media, sensors, images or speech, that is processed using various technologies including computer software, algorithms, machine learning , natural language processing , artificial intelligence , augmented intelligence and robotics . The increasing use of automated decision-making systems (ADMS) across a range of contexts presents many benefits and challenges to human society requiring consideration of the technical, legal, ethical, societal, educational, economic and health consequences.

Overview

There are different definitions of ADM based on the level of automation involved. Some definitions suggests ADM involves decisions made through purely technological means without human input, such as the EU's General Data Protection Regulation (Article 22). However, ADM technologies and applications can take many forms ranging from decision-support systems that make recommendations for human decision-makers to act on, sometimes known as augmented intelligence or 'shared decision-making', to fully automated decision-making processes that make decisions on behalf of individuals or organizations without human involvement. Models used in automated decision-making systems can be as simple as checklists and decision trees through to artificial intelligence and deep neural networks (DNN).

Since the 1950s computers have gone from being able to do basic processing to having the capacity to undertake complex, ambiguous and highly skilled tasks such as image and speech recognition, gameplay, scientific and medical analysis and inferencing across multiple data sources. ADM is now being increasingly deployed across all sectors of society and many diverse domains from entertainment to transport.

An ADM system (ADMS) may involve multiple decision points, data sets, and technologies (ADMT) and may sit within a larger administrative or technical system such as a criminal justice system or business process.

Data

Automated decision-making involves using data as input to be analyzed within a process, model, or algorithm or for learning and generating new models. ADM systems may use and connect a wide range of data types and sources depending on the goals and contexts of the system, for example, sensor data for self-driving cars and robotics, identity data for security systems, demographic and financial data for public administration, medical records in health, criminal records in law. This can sometimes involve vast amounts of data and computing power.

Data quality

The quality of the available data and its ability to be used in ADM systems is fundamental to the outcomes. It is often highly problematic for many reasons. Datasets are often highly variable; corporations or governments may control large-scale data, restricted for privacy or security reasons, incomplete, biased, limited in terms of time or coverage, measuring and describing terms in different ways, and many other issues.

For machines to learn from data, large corpora are often required, which can be challenging to obtain or compute; however, where available, they have provided significant breakthroughs, for example, in diagnosing chest X-rays.

ADM technologies

Automated decision-making technologies (ADMT) are software-coded digital tools that automate the translation of input data to output data, contributing to the function of automated decision-making systems. There are a wide range of technologies in use across ADM applications and systems.

ADMTs involving basic computational operations

Search (includes 1-2-1, 1-2-many, data matching/merge)

Matching (two different things)

Mathematical Calculation (formula)

ADMTs for assessment and grouping:

User profiling

Recommender systems

Clustering

Classification

Feature learning

Predictive analytics (includes forecasting)

ADMTs relating to space and flows:

Social network analysis (includes link prediction)

Mapping

Routing

ADMTs for processing of complex data formats

Image processing

Audio processing

Natural Language Processing (NLP)

Other ADMT

Business rules management systems

Time series analysis

Anomaly detection

Modelling/Simulation

Machine learning

Machine learning (ML) involves training computer programs through exposure to large data sets and examples to learn from experience and solve problems. Machine learning can be used to generate and analyse data as well as make algorithmic calculations and has been applied to image and speech recognition, translations, text, data and simulations. While machine learning has been around for some time, it is becoming increasingly powerful due to recent breakthroughs in training deep neural networks (DNNs), and dramatic increases in data storage capacity and computational power with GPU coprocessors and cloud computing.

Machine learning systems based on foundation models run on deep neural networks and use pattern matching to train a single huge system on large amounts of general data such as text and

images. Early models tended to start from scratch for each new problem however since the early 2020s many are able to be adapted to new problems. Examples of these technologies include Open AI's DALL-E (an image creation program) and their various GPT language models, and Google's PaLM language model program.

Applications

ADM is being used to replace or augment human decision-making by both public and private-sector organisations for a range of reasons including to help increase consistency, improve efficiency, reduce costs and enable new solutions to complex problems.

Debate

Research and development are underway into uses of technology to assess argument quality, assess argumentative essays and judge debates. Potential applications of these argument technologies span education and society. Scenarios to consider, in these regards, include those involving the assessment and evaluation of conversational , mathematical , scientific , interpretive , legal , and political argumentation and debate.

Law

In legal systems around the world, algorithmic tools such as risk assessment instruments (RAI), are being used to supplement or replace the human judgment of judges, civil servants and police officers in many contexts. In the United States RAI are being used to generate scores to predict the risk of recidivism in pre-trial detention and sentencing decisions, evaluate parole for prisoners and to predict "hot spots" for future crime. These scores may result in automatic effects or may be used to inform decisions made by officials within the justice system. In Canada ADM has been used since 2014 to automate certain activities conducted by immigration officials and to support the evaluation of some immigrant and visitor applications.

Economics

Automated decision-making systems are used in certain computer programs to create buy and sell orders related to specific financial transactions and automatically submit the orders in the international markets. Computer programs can automatically generate orders based on predefined set of rules using trading strategies which are based on technical analyses, advanced statistical and mathematical computations, or inputs from other electronic sources.

Business

Continuous auditing

Continuous auditing uses advanced analytical tools to automate auditing processes. It can be utilized in the private sector by business enterprises and in the public sector by governmental organizations and municipalities. As artificial intelligence and machine learning continue to advance, accountants and auditors may make use of increasingly sophisticated algorithms which make decisions such as those involving determining what is anomalous, whether to notify personnel, and how to prioritize those tasks assigned to personnel.

Media and entertainment

Digital media, entertainment platforms, and information services increasingly provide content to audiences via automated recommender systems based on demographic information, previous selections, collaborative filtering or content-based filtering. This includes music and video platforms, publishing, health information, product databases and search engines. Many recommender systems also provide some agency to users in accepting recommendations and incorporate data-driven algorithmic feedback loops based on the actions of the system user.

Large-scale machine learning language models and image creation programs being developed by companies such as OpenAI and Google in the 2020s have restricted access however they are likely to have widespread application in fields such as advertising, copywriting , stock imagery and graphic design as well as other fields such as journalism and law.

Advertising

Online advertising is closely integrated with many digital media platforms, websites and search engines and often involves automated delivery of display advertisements in diverse formats. 'Programmatic' online advertising involves automating the sale and delivery of digital advertising on websites and platforms via software rather than direct human decision-making. This is sometimes known as the waterfall model which involves a sequence of steps across various systems and players: publishers and data management platforms, user data, ad servers and their delivery data, inventory management systems, ad traders and ad exchanges. There are various issues with this system including lack of transparency for advertisers, unverifiable metrics, lack of control over ad venues, audience tracking and privacy concerns. Internet users who dislike ads have adopted counter measures such as ad blocking technologies which allow users to automatically filter unwanted advertising from websites and some internet applications. In 2017, 24% of Australian internet users had ad blockers.

Health

Deep learning AI image models are being used for reviewing x-rays and detecting the eye condition macular degeneration.

Social services

Governments have been implementing digital technologies to provide more efficient administration and social services since the early 2000s, often referred to as e-government . Many governments around the world are now using automated, algorithmic systems for profiling and targeting policies and services including algorithmic policing based on risks, surveillance sorting of people such as airport screening, providing services based on risk profiles in child protection, providing employment services and governing the unemployed. A significant application of ADM in social services relates to the use of predictive analytics – eg predictions of risks to children from abuse/neglect in child protection , predictions of recidivism or crime in policing and criminal justice, predictions of welfare/tax fraud in compliance systems, predictions of long term unemployment in employment services. Historically these systems were based on standard statistical analyses, however from the early 2000s machine learning has increasingly been developed and deployed. Key issues with the use of ADM in social services include bias, fairness, accountability and explainability which refers to transparency around the reasons for a decision and the ability to explain the basis on which a machine made a decision. For example Australia's federal social security delivery agency, Centrelink, developed and implemented an automated processes for detecting and collecting debt which led to many cases of wrongful debt collection in what became known as the RoboDebt scheme.

Transport and mobility

Connected and automated mobility (CAM) involves autonomous vehicles such as self-driving cars and other forms of transport which use automated decision-making systems to replace various aspects of human control of the vehicle. This can range from level 0 (complete human driving) to level 5 (completely autonomous). At level 5 the machine is able to make decisions to control the vehicle based on data models and geospatial mapping and real-time sensors and processing of the environment. Cars with levels 1 to 3 are already available on the market in 2021. In 2016 The German government established an 'Ethics Commission on Automated and Connected Driving' which recommended connected and automated vehicles (CAVs) be developed if the systems cause fewer accidents than human drivers (positive balance of risk). It also provided 20 ethical rules for the adaptation of automated and connected driving. In 2020 the European Commission strategy on CAMs recommended that they be adopted in Europe to reduce road fatalities and lower emissions however self-driving cars also raise many policy, security and legal issues in terms of liability and ethical decision-making in the case of accidents, as well as privacy issues. Issues of trust in autonomous vehicles and community concerns about their safety are key factors to be addressed if AVs are to be widely adopted.

Surveillance

Automated digital data collections via sensors, cameras, online transactions and social media have significantly expanded the scope, scale, and goals of surveillance practices and institutions in government and commercial sectors. As a result there has been a major shift from targeted

monitoring of suspects to the ability to monitor entire populations. The level of surveillance now possible as a result of automated data collection has been described as surveillance capitalism or surveillance economy to indicate the way digital media involves large-scale tracking and accumulation of data on every interaction.

Ethical and legal issues

There are many social, ethical and legal implications of automated decision-making systems. Concerns raised include lack of transparency and contestability of decisions, incursions on privacy and surveillance, exacerbating systemic bias and inequality due to data and algorithmic bias , intellectual property rights, the spread of misinformation via media platforms, administrative discrimination, risk and responsibility, unemployment and many others. As ADM becomes more ubiquitous there is greater need to address the ethical challenges to ensure good governance in information societies.

ADM systems are often based on machine learning and algorithms which are not easily able to be viewed or analysed, leading to concerns that they are 'black box' systems which are not transparent or accountable.

A report from Citizen Lab in Canada argues for a critical human rights analysis of the application of ADM in various areas to ensure the use of automated decision-making does not result in infringements on rights, including the rights to equality and non-discrimination; freedom of movement, expression, religion, and association; privacy rights and the rights to life, liberty, and security of the person.

Legislative responses to ADM include:

The European General Data Protection Regulation (GDPR) , introduced in 2016, is a regulation in EU law on data protection and privacy in the European Union (EU). Article 22(1) enshrines the right of data subjects not to be subject to decisions, which have legal or other significant effects, being based solely on automatic individual decision making. GDPR also includes some rules on the right to explanation however the exact scope and nature of these is currently subject to pending review by the Court of Justice of the European Union . These provisions were not first introduced in the GDPR, but have been present in a similar form across Europe since the Data Protection Directive in 1995, and the 1978 French law, the loi informatique et libertés [fr] . Similarly scoped and worded provisions with varying attached rights and obligations are present in the data protection laws of many other jurisdictions across the world, including Uganda , Morocco and the US state of Virginia .

Rights for the explanation of public sector automated decisions forming 'algorithmic treatment' under the French loi pour une République numérique .

Bias

ADM may incorporate algorithmic bias arising from:

Data sources, where data inputs are biased in their collection or selection

Technical design of the algorithm, for example where assumptions have been made about how a person will behave

Emergent bias, where the application of ADM in unanticipated circumstances creates a biased outcome

Explainability

Questions of biased or incorrect data or algorithms and concerns that some ADMs are black box technologies, closed to human scrutiny or interrogation, has led to what is referred to as the issue of explainability, or the right to an explanation of automated decisions and AI. This is also known as Explainable AI (XAI), or Interpretable AI, in which the results of the solution can be analysed and understood by humans. XAI algorithms are considered to follow three principles - transparency, interpretability and explainability.

Information asymmetry

Automated decision-making may increase the information asymmetry between individuals whose data feeds into the system and the platforms and decision-making systems capable of inferring information from that data. On the other hand it has been observed that in financial trading the information asymmetry between two artificial intelligent agents may be much less than between two human agents or between human and machine agents. A research validated Daniel Kahneman's theory on noisy decisions by human experts in finance. It demonstrates the inherent inconsistencies in human judgments, which consequently affect the outcomes of automated decisions made by AI decision-support systems.

Research fields

Many academic disciplines and fields are increasingly turning their attention to the development, application and implications of ADM including business, computer sciences, human computer interaction (HCI), law, public administration, and media and communications. The automation of media content and algorithmically driven news, video and other content via search systems and platforms is a major focus of academic research in media studies.

The ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT) was established in 2018 to study transparency and explainability in the context of socio-technical systems, many of which include ADM and AI.

Key research centres investigating ADM include:

Algorithm Watch, Germany

ARC Centre of Excellence for Automated Decision-Making and Society , Australia

Citizen Lab , Canada

Informatics Europe

See also

Automated decision support

Algorithmic bias

Decision-making software

Decision Management

Ethics of artificial intelligence

Government by algorithm

Machine learning

Recommender systems

References