Abstract: The application of Artificial Intelligence or AI in education has been the subject of academic research for more than 30 years. The field examines learning wherever it occurs, in traditional classrooms or at workplaces so to support formal education and lifelong learning. It combines interdisciplinary AI and learning sciences (such as education, psychology, neuroscience, linguistics, sociology and anthropology) in order to facilitate the development of effective adaptive learning environments and various flexible, inclusive tools. Nowadays, there are several new challenges in the field of education technology in the era of smart phones, tablets, cloud computing, Big Data, etc., whose current research questions focus on concepts such as ICT-enabled personalized learning, mobile learning, educational games, collaborative learning on social media, MOOCs, augmented reality application in education and so on. Therefore, to meet these new challenges in education, several fields of research using AI have emerged over time to improve teaching and learning using digital technologies. Moreover, each field of research is distinguished by its own vision and methodologies. In this article, the authors present a state of the art finding in the fields of research of Artificial Intelligence in Education or AIED, Educational Data Mining or EDM and Learning Analytics or LA. We discuss their historical elements, definition attempts, objectives, adopted methodologies, application examples and challenges.

Keywords: personalized learning; mobile learning; educational; collaborative learning on social media MOOCs; AIED; EDM; LA

1. Introduction

The application of AI in education has been the subject of academic research for more than 30 years. The field examines learning wherever it occurs, in traditional classrooms or at workplaces so to support formal education and lifelong learning. It combines interdisciplinary AI and learning sciences (such as education, psychology, neuroscience, linguistics, sociology and anthropology) in order to facilitate the development of effective adaptive learning environments and various flexible, inclusive tools. It is personalized and attractive for teaching and learning [1]. The same goes with the AIED, which focuses on issues related to the theories of human learning and AI application in effective learning environments, as well as theories of teaching and AI application to effective educational systems. It is clear that in many cases, there is a fuzzy boundary between learning environments and teaching systems [2].

2. Artificial Intelligence in Education (AIED)

2.1. The Objectives of the AIED

The scientific goal of AIED is to define specific and explicit forms of knowledge about education, including several psychological and social aspects which often remain implicit [3]. In addition to driving many “smart” technologies, AIED is intended as a
powerful way to explore in detail what is sometimes called the “black box of learning.” This helps to better understand how learning actually occurs; to analyze the influence of the socio-economic factors of the learner [4], the physical context and the technology [5] or to study the nature of knowledge and its representation. Determining the most appropriate way to learn and the most effective teaching interaction styles helps a learner in his/her learning, especially when it is used in the right moments. To identifying the misconceptions that learners have about the learning object, AIED effectively involves two complementary components: developing AI-based tools to support learning and using these tools to help understand learning. For example, by modeling how learners solve an arithmetic problem and by identifying misconceptions previously unknown to educators, researchers and teachers can learn much more about the learning process itself, and these understandings could then be applied to classical classroom practices [6]. In addition, they could be integrated into the development of AIED tools [1].

2.2. The Strategy of the AIED

Researchers in AIED are paying increasing attention to the emotional [7], social [3] and intellectual aspects of learning with very active research conducted in the study of collaboration [8], metacognition [9], self-regulation, motivation and emotions [10]. This research is motivated by educational problems and focuses as much on research as on technological development. Research based on theory is supported by a systematic empirical evaluation that informs the further development of the theory. The AIED community is actively exploring ways in which learning and education can take advantages of new and advanced technologies, including advances in AI [11].

2.3. Example of AIED Tools: Intelligent Tutoring Systems

Intelligent tutoring systems or ITS is one of the most common applications of AI in education, or at least it is probably the oldest. An ITS generally offers step-by-step tutorials on topics in well-defined and structured subjects such as mathematics or physics, which are customized for each learner. That is, it relies on specialized knowledge about the subject and a pedagogical approach. In response to the misconceptions and correctness of each learner, the system determines step by step an optimal path through the support and learning activities. As the learner progresses, the system automatically adjusts the level of difficulty and provides hints or tips that all aim to ensure that the learner is able to learn the given subject effectively [12]. Some ITSs allow learners to control their own learning to help them develop self-regulatory skills; others use instructional strategies to regulate the progression of learning to support the learner [1]. ITSs are based on models that represent knowledge specific to teaching and learning. In general, there are three types of knowledge. Firstly, knowledge about the subject to be learned is represented in what is so-called a domain model. Secondly, knowledge about effective teaching approaches is represented in a pedagogical model. Thirdly, knowledge about the learner is represented in a learner model. From these three models, algorithms can adapt a sequence of learning activities to each learner [13]. Instead of models, many recent ITSs use machine learning techniques, self-learning algorithms based on large data sets and neural networks to enable them to make appropriate content which then is provided to the learner. However, with this approach, it may be difficult to explain the rationale for these decisions [1] (Figure 1).

2.3.1. The Domain Model

A domain model represents the knowledge that ITS aims to help learners acquire. This may include, for example, knowledge of mathematical procedures, genetic heritage, or causes of the First World War [10]. In fact, mathematics for elementary and high school students has dominated ITSs over the years. Physics and computer science are also fruits within reach of ITSs because they are, at least at basic levels, well-structured and clearly defined [14].
2.3.2. The Educational Model

The pedagogical model represents knowledge of effective teaching and learning approaches that have been obtained from pedagogical experts and research in the learning sciences [14]. The pedagogical knowledge that has been represented in many ITSs includes knowledge of pedagogical approaches [15], proximal developmental area, interlaced practice, cognitive load and formative return. For example, a teaching model using the Vygotsky Proximal Development Zone ensures that the activities provided by the system to the learner are neither too easy nor too stimulating. A model implementing an individualized formative return ensures that the return is provided to the learner whenever it is possible to give support.

Figure 1. Simplified Schema of ITS based on a model.
2.3.3. The Model of the Learner

What distinguishes AI-based ITSs is that they also include a learner model; that is, a representation of the learner’s state of knowledge. In fact, many ITSs incorporate a wide range of knowledge about the learner such as their interactions, the material that are challenging to the learner, their misconceptions, and their emotional states when using the system. This information can be used to inform the progress of the learning process and therefore determine the support that will be given to the learner. When most ITSs go much further, the knowledge stored on each learner is supplemented by the knowledge of all learners who have already used the system. From the data of all learners, therefore, the system learns to predict which pedagogical approach and field are appropriate for a particular learner. It is the learner model that allows ITSs to be adaptive, and machine learning makes this adaptive process more efficient [16].

3. Educational Data Mining (EDM)

EDM is an emerging field linked to several established research areas, including e-learning, adaptive hypermedia, intelligent tutoring systems, online exploration, data mining and so on. The application of data mining in education systems has specific requirements that are not present in other areas, mainly the need to take into account the pedagogical aspects of the learner and the system. Although the exploration of educational data is a very recent field of research, a large number of contributions published in journals, international congresses, specific workshops and works in progress [17] show that is a promising new field. EDM is concerned with developing, researching and applying computer-based methods to detect schemas of large educational data collections [18]. These patterns would, otherwise, be difficult or impossible to analyze directly because of the huge amount of data in which they exist. Data of interest is not limited to learner interactions with an educational system (e.g., navigation behavior, questionnaire entry and interactive exercises), but it may also include data from collaborating learners (textual dialogue, for example), administrative data (school, etc.), and demographic data (such as gender, age, school results). Data on learner states (motivation, emotional states, for example) has also been taken into account, which can be deduced from physiological sensors (facial expression, seat posture and perspiration, for example). EDM uses methods and tools from the broader field of data mining [8,19], and a subdomain of computer science and AI that has been used for purposes as diverse as credit card fraud detection, genetic sequence analysis in bioinformatics, or analysis of customer buying behavior [20].

3.1. Attempts to Define EDM

EDM is defined as the area of scientific inquiry focused on the development of methods for discovering types of data uniquely derived from educational contexts, and for using these methods to better understand learners and the context learning [8,15]. In other words, EDM is about converting raw data from educational systems into useful information that can be used to inform design decisions and answer research questions [21].

3.2. The Main Approaches in EDM

Data mining, in general, encompasses a wide range of search techniques that include more traditional options such as database queries and simple automatic logging, as well as more recent developments in machine learning and linguistic technology [21]. EDM methods often differ from methods in the broader literature on data mining, explicitly exploiting the multiple, hierarchical, significant levels of educational data. Psychometric methods are often incorporated into machine learning and data mining methods to achieve this goal [22]. As a result of this, there is a wide variety of common popular methods in educational data mining (Table 1). These methods fall into the following general categories: forecasting, grouping, relationship exploration, discovery with models and data distillation for human judgment. The first three categories are widely recognized as universal for all types of data mining (although they are in some cases under different names). The fourth
and fifth categories gain special importance in the exploration of educational data, and so on [22].

A large number of EDM applications have been used, so we have to pay particular attention to four areas of application. One of the main areas of application is the improvement of learner models, which is providing detailed information on learner characteristics or states such as knowledge, motivation, meta-cognition and attitude. Modelling individual differences between learners to allow software to address these differences is a key theme in educational software research. A second key application area is to discover or improve models of the domain knowledge structure. In EDM, methods have been created for rapid discovery of specific domain models directly from data. A third key application area is to study the educational support provided by learning software. Modern educational software offers various types of educational support to learners. Finding the most effective educational support is a key area of interest in EDM. The decomposition of learning, a type of relationship exploration, adapts exponential learning curves to performance data so to link learner success to the quantity of each type of instructional medium that a learner has received (with a weight for each type of support) [23]. The weights indicate the effectiveness of each type of pedagogical support to improve learning. A fourth key application area of the EDM is scientific discovery about learning and learners. It therefore takes many forms. The EDM to answer questions in one of the three areas previously discussed (such as learner models, domain models and pedagogical support) may have broader scientific benefits. For example, the study of pedagogical support may have the long-term potential to enrich the theories of learning [24].

Table 1. The main categories of analyzes in EDM.

<table>
<thead>
<tr>
<th>Category of Method</th>
<th>Goals of Method</th>
<th>Key Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction</td>
<td>Develop a model which can infer a single aspect of the data (predicted variable) from some combination of other aspects of the data (predictor variables)</td>
<td>Detecting student behaviors (e.g., gaming the system, off-task behavior, slipping); developing domain models; predicting and understanding student educational outcomes.</td>
</tr>
<tr>
<td>Clustering</td>
<td>Find data points that naturally group together, slipping the full data set into a set of categories.</td>
<td>Discovery of new student behavior patterns; investigating similarities and differences between schools</td>
</tr>
<tr>
<td>Relationship mining</td>
<td>Discover relationships between variables</td>
<td>Discovery of curricular associations in course sequences; Discovering which pedagogical strategies lead to more effective/robust learning</td>
</tr>
<tr>
<td>Discovery with models</td>
<td>A model of a phenomenon developed with prediction, clustering or knowledge engineering, is used as a component in further prediction or relationship mining</td>
<td>Discovery of relationships between student behaviors, and student characteristics or contextual variables; Analysis of research questions across wide variety of contexts.</td>
</tr>
<tr>
<td>Distillation of data for Human judgment</td>
<td>Data is distilled to enable a human to quickly identify or classify features of the data.</td>
<td>Human identification of patterns in student learning, behavior, or collaboration; Labeling data for use in later development of prediction model</td>
</tr>
</tbody>
</table>

3.3. Examples of Applications of EDM

A typical feature of educational data is its non-independence. That is, when we collect data from educational discussions and when we want to rank whether the contributions are on a topic or not, we must consider that the contributions are not statistically independent of each other because many contributions may come from the same learner or discussion. This could be detrimental to the calculation of models (standard machine learning schemes usually include the built-in assumption of independent training examples) as well as model
validation (for example, cross-validation could lead to biased results when the training and the set of tests are not independent).

In addition, the results of research in EDM are generally obtained in the narrow context of specific research projects and educational contexts (such as a particular school). The question is how general these results are, for instance, if the same learner model parameters can also be used with other learner populations or if a predictive model is always reliable in a different context. Therefore, there is a growing need for replication studies to test broader generalizations. As a result, EDM researchers are increasingly interested in open data repositories and standard data formats to promote the exchange of data and models [18].

EDM is a young field of research, and it is necessary to initiate more specialized and oriented professional training in order to achieve a level of success similar to that of other fields such as the extraction of medical data, extraction of e-commerce data, and so on [25].


4.1. Definition of Learning Analytics

Learning Analysis is an emerging field concerned with analyzing the vast data of learners from the environments supported by technology to inform the theory and practice of education. Since its creation, it has adopted a multidisciplinary approach integrating learning studies and technological capabilities [2]. According to SoLAR [2], learning analysis is the measurement, collection, analysis and communication of data about learners and their contexts for understanding and optimizing purposes, learning and the environments in which it occurs [23].

4.2. Objectives of LA

A multitude of factors motivated interest in learning analysis. The general trend towards increased accountability at all levels of education is a motivating factor for the increased interest in learning analysis. Educational institutions are increasingly eager to report what their learners are learning and where there is even greater pressure for e-learning as these courses now have separated accreditation standards. Learning analysis not only documents learners’ performance, but it also provides tools that encourage the types of continuous improvement sought by accreditors. In addition, there are other goals that the analysis of learning aims to achieve. These may include predicting learner performance, suggesting learners relevant educational resources, increasing learner awareness and reflection, detecting unwanted learning behaviors, and detecting emotional states (e.g., boredom, frustration). It should be remembered that teachers were essentially inspired by their intuition to know when learners are struggling, when to propose relevant educational resources or how to encourage them to reflect on their learning. These intuitions will not disappear with the advent of the analytics of learning, of the actions that flow from them. The analytics of learning promises to make these intuitions and the resulting actions more data-driven and easier to detect.

4.3. Research Methodology in Learning Analytics

Research in learning analysis should clearly indicate how the proposed work offers new and relevant analytical methods (e.g., methods of calculation, representation, statistics and visualization) or should improve our understanding of the value of existing analytical methods in the literature. Therefore, research on learning analysis may vary in technical contribution, but the link with learning must be present. There are two cases to understand learning and teaching practices. Research with strong technical input does not need to include a study in an applied context, but it must at least explain how the properties of the technical contribution are relevant to understanding or managing learning in practice. Research with a theoretical or practical contribution to learning does not necessarily have to advance the technical state of the analysis, but it must at least examine or evaluate whether and how the advantages of the chosen analytical methods generate the relevant aspects of the analytical data, so to allow the main contribution [3].
No particular theory or design of learning should be favored a priori: individuals, small groups and/or larger collectives can be agents of learning; and learning may include the acquisition of knowledge or skills, the definition of an inter-subjective meaning, or changes in identity and participation in a community among other processes [24]. In addition, learning can be conceived as taking place simultaneously with all these granularities and involving all these epistemological processes. Research that analyzes learning processes across multiple granularities and that provides multiple methodological and theoretical orientations is particularly appropriate for understanding learning as a complex phenomenon. Regardless of how learning is conceptualized, the goal is to encourage learning analytics researchers to make their design explicit and to think about the links between their analytic approach and their understanding of learning.

Therefore, Learning Analysis is a rapidly expanding field of educational technology research. It has strong roots in a variety of areas, including business intelligence, web analytics, data mining and referral systems. Its close links to these areas mean that researchers and practitioners have approached it from different angles, and they must now work together to identify not only the goals that can be achieved through learning analysis but also the measures to achieve these objectives [25].

### 4.4. Types of Data Used in Learning Analytics

Educational institutions have a wealth of data that can help improve learners’ performance and increase their motivation. Hence, we present some data that teachers are likely to have at their fingertips and that are amenable to some basic data analysis. In the table below (Table 2), the types of data available in an LMS are presented. The first column includes data automatically generated by the LMS. The second column is an example of the types of data that can be generated by the instructor, much of which can be stored in the LMS.

<table>
<thead>
<tr>
<th>Data Generated by LMS</th>
<th>Data Generated by Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of times resource accessed</td>
<td>Grades on discussion forum</td>
</tr>
<tr>
<td>Date and time of access</td>
<td>Grades on Assignment</td>
</tr>
<tr>
<td>Number of discussion posts generated</td>
<td>Grades on tests</td>
</tr>
<tr>
<td>Number of discussion posts read</td>
<td>Final Grades</td>
</tr>
<tr>
<td>Types of resource accessed</td>
<td>Number (and type) of questions asked in a discussion forum.</td>
</tr>
<tr>
<td></td>
<td>Number of Emails sent to instructor</td>
</tr>
</tbody>
</table>

### 4.5. LA and EDM: Similarities and Differences

Educational Data Mining (EDM) and Learning Analysis (LA) are relatively new and promising areas of research aimed at improving educational experiences by helping stakeholders (trainers, learners, administrators and researchers) make better decisions using data. Their growth has been stimulated by the increased capacity of computers to store and analyze large amounts of data and the availability of statistical methods and techniques, machine learning and data mining. Online environments are an extremely important area of application. On the one hand, they continuously generate data from many events such as reading files or participating in forums, with different formats and levels of hierarchy. Similarly, online courses have higher drop-out rates than traditional courses. EDM and LA are mainly used to monitor learners and groups (to identify students who are likely to drop out or fail, or who do not contribute enough to collaborative activities), suggest changes in course structure and experiences tailor-made learning (recommending materials based on needs, motivations and skills, for example). There is a wide variety of methods and techniques adapted from other disciplines or specifically designed to analyze educational data. There are many similarities between the two areas of research, such as objectives, methodologies and techniques. However, there are several differences mainly due to their
origins and tendencies. The coexistence of their respective scientific communities creates competition with positive effects for society.

Despite the high expectations and the amount of work on EDM and LA, their application in educational environments still faces significant challenges such as a lack of a data-driven culture and fast, comprehensive, easy-to-use tools to understand who could be integrated into the most popular learning management systems.

In the age of Big Data, the combination of the current capacity for capturing, storing, managing and processing data in a timely manner, and data from e-learning environments provides researchers with AL opportunities to better explore the learning processes of students and effective ways to improve them. An important application is that of MOOCs, where the data of thousands of learners can be used to redefine the courses of future learners by using navigation and the use of tools. A much more ambitious approach is to develop adaptive MOOCs in which courses are automatically customized according to the student profile (needs, goals, background, country, learning style, etc.) and performance. This is a relatively new subject of research that currently attracts the attention of researchers and companies.

The similarities between EDM and LA suggest many areas of overlapping research. In addition, organizational deployment of EDM and LA requires similar data sets and research skills. However, these two communities have different roots, so it is important to note some distinctions. Table 3 shows some of the main differences between the two communities. It is sufficient to note that these distinctions are meant to represent the major trends in both communities; many researchers in electro-erosion are conducting research that could be placed on the local knowledge of each of these distinctions, and many researchers in these regions are conducting research that could be placed on these distinctions in SHS. By identifying these distinctions, we hope to recognize places where both communities can learn from each other, rather than defining communities exclusively. Indeed, communities that develop organically like these two communities will not have a rigid boundary between the work that appears in both communities. A key distinction lies in the type of discovery prioritized. In both communities, research can be found that uses automated discoveries and research that harness human judgment through visualization and other methods. However, EDM puts a lot more emphasis on automated discovery, and LA relies much more on human judgment. Even if researchers combine these two directions, this preference can be seen; that is, EDM research, which is often based on human judgment, provides labels for classification, while LA research using auto-discovery often informs humans who make the final decisions. This difference is associated with another one between the two communities: the type of adaptation and customization generally supported by both communities. In addition to the emphasis on automated discovery in EDM, whose models are more often used as a basis for automated adaptation carried out by a computer system such as a smart tutoring system. In contrast, LA models are more often designed to inform and empower instructors and learners. A third and important difference is the distinction between holistic and reductionist frameworks. It is much more common in EDM to see research that reduces phenomena to components and that analyzes individual components and their relationships. The "discovery with models" paradigm for EDM research presented in [23] is a clear example of this paradigm. On the other hand, LA researchers generally emphasize the need to understand systems as sets in all their complexity. The debate between reductionist and holistic paradigms has often paralyzed discussions between educational researchers of different “camps”. Encouraging discussions between EDM and LA researchers is an essential way to prevent this common split from reducing what EDM and LA researchers can learn from each other.

The methodologies used in EDM and LA come from a number of sources, but the two most important sources of inspiration are methods of data mining and analysis in general, as well as psychometric and educational measurement. In many cases, the specific characteristics of education data have led to different methods generally playing a
greater role in EDM and LA than in data mining, or leading to adaptations to methods of existing psychometrics:

- Prediction methods,
- Discovery of the structure,
- Mining relationship,
- Distillation of data for human judgment,
- Discovery with models, and
- Tools for conducting EDM/LA methods.

Table 3. A brief comparison of the two fields.

<table>
<thead>
<tr>
<th></th>
<th>LA</th>
<th>EDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>Leveraging human judgment is key; automated discovery is a tool to accomplish this goal</td>
<td>Automated discovery is key; leveraging human judgment is a tool to accomplish this goal</td>
</tr>
<tr>
<td>Reduction &amp; holism</td>
<td>Stronger emphasis on understanding systems as wholes, in their full complexity.</td>
<td>Stronger emphasis on reducing to components and analyzing individual components between them</td>
</tr>
<tr>
<td>Origins</td>
<td>LA has stronger origins in semantic web, “intelligent curriculum”, outcome prediction, and systemic interventions.</td>
<td>LED has strong origins in educational software and student modeling, with a significant community in predicting course outcomes.</td>
</tr>
<tr>
<td>Adaptation and personalization</td>
<td>Greater focus on informing and empowering instructors and learners Social network analysis, sentiment analysis, influence analytics, discourse analysis, learner success prediction, concept analysis, sensemaking models</td>
<td>Greater focus on automated adaption (e.g., by te computer with no human in the loop)</td>
</tr>
<tr>
<td>Techniques and methods</td>
<td>Classification, clustering, Bayesian modeling, relationship mining, visualization</td>
<td></td>
</tr>
</tbody>
</table>

5. Conclusions

Through this paper, we carried out a characterization of research work on the application of AI in education. The oldest scientific community, AIED, stands out above all for its research on the development of intelligent systems based on the modeling of elements of the learning context such as knowledge, teachers, learners, etc. Previously, these models were based on knowledge representation techniques, as well as the profiling of teachers and learners through their respective attributes. After that, we now have a large amount of varied data on education available with the advent of Big Data. This paved the way for new research work including the field of EDM. Here, we seek to develop systems that draw their intelligence from the exploitation of educational data through the extraction of relevant information about the learning context and their uses by these systems. In the same context of Big Data, there has also been the emergence of research work on the issues of the collection, measurement, analysis and communication of educational data. Thus, we are in the context of developing decision support tools in a learning context. In other words, it is the stakeholders, including institutions, educational managers and learners who constitute the direct consumers of the information provided by the LAs. Moreover, it also appears that these different research works are complementary for the improvement of learning and teaching through ICTs.

Author Contributions: Conceptualization, O.H.; methodology, O.H.; software, O.H, J.L. and M.H.A.H.; validation, N.-E.E.F.; formal analysis, O.H. and M.H.A.H.; investigation, O.H. and N.-E.E.F.; writing—original draft preparation O.H. and M.H.A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available from the authors upon request.

Acknowledgments: The authors would like to thank the individuals who generously shared their time and experience for supporting the purpose of this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

