

An Analysis of Big Data Analytics in Relation to Artificial Intelligence and Business Intelligence

Divya Sai Jaladi^{1*}, Sandeep Vutla²

¹Senior Lead Application Developer, SCDMV, 10311 Wilson Boulevard, Blythewood, SC 29016, UNITED STATES

²Assistant Vice President, Senior-Data Engineer, Chubb, 202 Halls Mill Rd, Whitehouse Station, NJ 08889, UNITED STATES

*Corresponding Author: divyasaj26@gmail.com

Abstract

Through the integration of analytical tools with operational and historical data, business intelligence systems provide crucial and competitive information to those responsible for corporate planning and decision-making. In order to help managers, understand how their organisation stacks up against competitors, business intelligence (BI) strives to improve the timeliness and quality of data. Business intelligence tools and technologies allow for the analysis of many market variables, including changes in market share, customer preferences, company capabilities, consumer behaviour and spending patterns, and market situations. Moreover, managers and analysts may use business intelligence to figure out which adjustments are going to be the most flexible in response to changing patterns. Data mining refers to the process of non-trivial discovery of hidden, useful, and potentially helpful information inside datasets. Technical methods include things like data summarisation, clustering, learning classification rules, finding dependency networks, analysing changes, and detecting anomalies. A more robust business intelligence environment is now available than in the past because to developments in data cleansing, improvements in hardware and software capabilities, the rise of web architecture, and the data warehouse as a repository. An attempt at a blueprint for building a BI system is made in this paper. Searches for and investigations into security holes have made use of AI. Acts of manipulation and mobility Robots powered by artificial intelligence are able to see and navigate in a static setting with relative ease.

Keywords: AI; Analytics; Big Data; Business Intelligence; Data Mining

Introduction

Making a living, providing for one's family, and contributing to a better world are all goals of business. Documentation of business activities, whether on paper or in digital form, is the first step in turning these records into data. Therefore, there is more data accessible on customer responses and the sector as a whole. Using specialised tools and techniques, this data may be analysed and mined to uncover insight and patterns that represent the company's operations. Incorporating these recommendations into the business could help it grow and better satisfy customer needs [1]. This repeats indefinitely. Organisations across all sectors may benefit from business intelligence, which includes methods and tools for gathering, analysing, and visualising data to help with decision-making at the executive level. Data mining is the process of extracting useful information from large datasets by use of statistical and machine learning techniques. Clustering, decision trees,

regression, and artificial neural networks additional data mining methods are included in this edition. Big data, text mining, and web mining are also covered in an easy-to-understand manner.

In order to swiftly adjust their future initiatives, each company must continually assess their business operations and advancement.

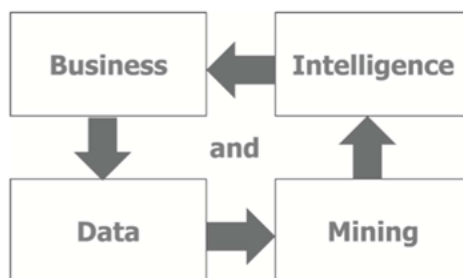


Figure 1. Business intelligence and data mining cycle.

Business intelligence include monitoring the market, competitors, vendors, and consumers. Key performance indicators (KPIs) or key result areas typically inform executives' metric selection processes [2]. Each executive needs their own personalised report to get the information they want. You may transform these reports into personalised dashboards that provide information in a way that is easy to comprehend and consume. A wide range of information technology (IT) solutions come together to produce business intelligence, which include tools for collecting, analysing, and reporting data to users on an organization's performance and its environment. Among the most important investment options are these IT solutions.

Think of a retail chain that has locations all over the globe and sells a wide range of products and services online as well. It compiles data on expenditures, sales, and purchases from different locations and time periods. Among other things, this data could be useful for identifying products that are selling well, products that are selling well in certain regions or during certain seasons, and customer groups that are seeing rapid growth. Additionally, it may assist you in considering which products go well together, which customers purchase certain items, and so forth. Better advertising campaigns, product bundling, and store layouts might result from applying this information [3]. Data mining is the practice of discovering and extracting useful information or decision-making expertise from massive datasets for applications such as decision support, estimation, forecasting, and prediction. Despite its size, the data is sometimes useless since it cannot be used directly; yet, the data's concealed information can be useful. A pattern is any kind of model or design that helps to explain or make sense of anything. Identifying patterns helps to establish relationships between elements that at first glance appear to have no connection. In order to make trends more easily understandable, patterns might be useful in reducing visual noise. Consistent patterns, like the sun rising in the east every day, may be as unyielding as scientific truths.

The term "artificial intelligence" (AI) describes machine intelligence. The term "artificial intelligence" (AI) refers to a method of programming computers to mimic human intelligence by assembling a set of algorithms that, when executed in parallel, may complete activities normally performed by people. The development of smart computers is facilitated by machine learning, a

branch of AI. Deep learning is a subfield of ML that represents data abstraction using well-established model architectures. Deep learning is a data-processing model that attempts to mimic the way the human brain works by generating patterns, minimisation (where possible), and accurate output. This study details the research concerns, AI approaches, applications, and the resources (hardware and software) that were used. In recent years, a wide range of fields have begun to employ ML, a subset of AI, to automate formerly manual and laborious tasks related to complex decision-making and problem-solving. Machine learning (ML) is an umbrella term for a set of methods that use past problems to train new computers to solve similar ones [4]. Artificial Neural Networks (ANNs) are the most well-known of the several methods that are accessible to begin with, it was an effort to mimic human learning abilities, with inspiration drawn from biological neural networks found in the brain. Natural language processing, genetic algorithms, inductive learning, case-based reasoning, and other approaches are also available.

A mid-century effort at NCSU's agricultural department laid the groundwork for what would later become known as the Statistical Analysis System (SAS), and it was this work that ushered in the modern digital age of data science [5]. Better and more trustworthy results than those obtained by business analysis were the initial motivation for introducing data science to the industry. Data processing, predictive modelling, and visualisation were among the sought-after skill sets for analysts working in data science. Right now, the languages that people use to process data the most are Python and R. Conversely, data processing and analysis may one day make use of Google's Go programming language. There is a plethora of tools, technology, and resources, which is causing data science to advance tremendously. It takes on a range of practical problems and offers potential remedies.

Review of Literature

Due to its roots in data science, business intelligence (BI) has grown rapidly in recent years, becoming a popular topic in DSS both academics and businesses alike. In order to help organisations make better decisions, it is possible to define it as an architecture, tool, technology, or system that saves and retrieves data, processes it using analytical tools, makes reporting and querying easier, and provides information and knowledge. Data science frameworks include BI, although the term has a far larger meaning. By taking massive amounts of data, processing it, and providing a high-level collection of reports that condense the core of that data into the foundation of business operations, business intelligence (BI) helps management make basic everyday business choices by acquiring high-quality and relevant information on the topic matter being studied. Business intelligence (BI) is defined by Cui et al. as a strategy and approach for improving company performance via providing strong support to senior decision-makers and making actionable information readily available to them. BI tools are a kind of technology that boosts the efficiency of business operations by making better use of corporate data and making that data more valuable overall. Zeng et al. define BI as "the process of gathering, processing, and disseminating information to reduce uncertainty in the making of all strategic choices." Business intelligence (BI) is defined as "applications and technology used to acquire, offer access to analysed data and information about a company to enable them to make more informed business choices" in the context of management. In contrast to its forerunner, "decision support," business intelligence (BI)

is a strategic tool for planning and performance monitoring, as opposed to Purdy operational choices, as stated by Van Drune [6].

Business intelligence (BI)

BI tools, according to Cui et al., have developed from EIS. When it comes to delivering information, Decision Support Systems (DSS) are head and shoulders above the competition. They can back up OLAP techniques including query, reporting, ad hoc analysis, and multidimensional analysis, among others. Corporations started pouring money into these kinds of intelligence systems because of their versatility. Still, business intelligence (BI) plans must to be well-defined and included into organisations' overall IT strategies [7]. Enterprise BI is a way to introduce new efficiencies across business sectors and synergies to company operations, whereas business intelligence (BI) is meant to describe an organization's ability to understand and use data for its advantage. By providing "one version of the truth," business intelligence helps companies maintain data consistency and harmonisation across all divisions. To guarantee data consistency across several applications in a complicated organisation, three main requirements must be satisfied, according to Arents:

Revisit: Make sure that all other applications are in sync with the system data.

Data accuracy requires that any information gathered from any other source be incorporated into the data set;

Users should be able to actively utilise the system as a decision-making tool if they have confidence in the data's timeliness and accuracy. Globalisation, deregulation, mergers & acquisitions, and technological innovation are only a few examples of the rapidly evolving economic dynamics that have forced companies to reevaluate their strategies [8]. Business intelligence (BI) is essential in today's competitive market because it helps with decision-making and boosts competitiveness by connecting corporate strategy with IT. In order to solve ever-more-complicated business challenges, BI technology is always developing and improving. Data mining, online analytical processing (OLAP), and data warehousing (DW) are the three most popular BI enabling technologies that have emerged (DM). By delivering timely, reliable, and relevant data, business intelligence (BI) systems aim to help people make "better" business decisions. Businesses that want to stay ahead of the competition collect business information (BI) to help them understand their environment and gain an advantage. BI may even be considered a core competency in certain companies.

Machine Learning

Tech for machine learning has been around for a while. The data-driven approach was superseded by machine learning in 1990. The focus shifted to information retrieval and natural language search. A renaissance of the neural network occurred in 2005, after it had been tested for the first time in 1957 for neural network computers [9]. There has been a lot of success and some failure with Machine Learning, but it has the potential to become widely used in the near future (2 to 5 years). In order to maintain parity with the machine's growth advancements in the field of education, as well as in the infrastructure and technological capacities that support it, are becoming increasingly important.

Deep Learning

In 1965, Deep Learning was created by Alexey Grigoryevich Ivakhnenko and Valentin Grigorievich Lapa. They used models that were statically analysed using complex equations and polynomial functions. It wasn't until 1995 that a method for finding and mapping comparable data was developed. Around 1997, recurrent neural networks were equipped with long short-term memory [10]. As high-speed processors were introduced in the late 1990s, eventually quadrupling computing speed by 1000, graphics processing units (GPUs) became more efficient at handling picture processing. The early 2000s saw the implementation of several pre-training layers and enhancements to lengthy short-term memory. Computers were able to run convolutional neural networks in 2011 without pre-training each layer, because to the improved speed of GPUs. Nowadays, massive volumes of data are handled via Deep Learning. New, more complex ideas are appearing as AI and Deep Learning continue to develop [11].

Artificial Intelligence (AI)

It is getting increasingly challenging for managers to completely understand the business sector due to the rapid change occurring within it and the increasing complexity of firm operations. Companies have had to reevaluate their approaches due to technology advancements, contestability, liberalisation, acquisitions, and mergers. Many major companies have used Business Intelligence (BI) methods to better understand and manage their business operations in order to obtain an edge over their competitors [12]. A primary goal of business intelligence (BI) is to help managers have a better grasp of the competitive landscape in which their organisation operates by improving the timeliness and quality of information. Through the use of business intelligence (BI) tools and technologies, companies may examine market conditions, client preferences, changing consumer behaviour and spending patterns, corporate capabilities, and trends in market share. Managers and analysts may use it to find out which changes will react best to new trends. It has developed into a standard for data analysis and decision-making support systems. The Ferranti Mark 1 began developing and executing AI algorithms in 1951. The field of artificial intelligence was established in 1956 during a summer workshop at Dartmouth College. Hardware resources for calculations were experiencing some issues at the moment. Businesses and the government invested billions of dollars in artificial intelligence development in the 1980s, in response to a request from the government. From 2000 to 2010, the area of artificial intelligence had a surge thanks to the money and attention shown to progress the topic. Machine learning emerged as a viable strategy for addressing a wide range of societal and industrial problems with the advent of powerful computing hardware.

The Function of AI in Analytics of Data

Data analytics and artificial intelligence are hot topics in academia throughout the world. Millions of money are being poured by developing nations into their pursuit of international fame. Nevertheless, very few allocate even a small portion of their resources to unclassified AI research. There are several industries that can benefit from artificial intelligence. Connecting with operational systems to import data into the warehouse is the responsibility of the acquisition component, which serves as the backbone of the data warehousing system. Databases such as Oracle, DB2, Informix, SQL Server, SAP R/3, and others are used to store data after it has been

input or processed through an OLTP-based daily business process. Before operational databases and external sources may contribute data to the data warehouse, the data must undergo the following processing steps:

Cleansing and Extraction

Data extraction involves collecting information from a variety of sources, one of which being operational systems. Then, the selected pollutants. As a last step, data purification involves checking and correcting extracted data for missing, inconsistent, or incorrect information. At this point, you may make use of triggers, error reporting, and corrective procedures [6].

Change

When data is transformed, it is transformed into a standard format and mapped to the warehouse schema using business rules. We aggregated and imputed attributes.

Import

Data loading is used to put the cleansed data into the data warehouse.

Methods for Machine Learning

To solve problems in a methodical way, decision trees use a hierarchical structure with variables or decision nodes. They come in helpful when evaluating sets of unique traits and qualities. One supervised learning model that has proven useful in regression and classification is the support vector machine analyses information [7]. You may use them to classify one variable as opposed to another, and the relationship between the variables doesn't even need to be linear. Feature independence and the Bayes theorem provide the basis of probabilistic classifiers, which in turn form the basis of the Naive Bayes classifier. The aggregate conditional probabilities of different attributes are computed by them [8].

Methods for Deep Learning

To calculate the weights of a network, gradients are generated using the backpropagation method. Two types of learning are available: batch and stochastic. As a deep learning training approach, dropout entails randomly removing units. Minimising the number of parameters is its main objective. By using the skip-gram approach, we may compare two terms in our vocabulary only if they are used in the same context. It randomly selects a word from the surrounding text when given the word in the middle of a sentence [9]. The probability of utilising a word in our vocabulary that is close by will be determined by the network.

Data Marts

Localised data warehouses, or data marts, are smaller versions of larger data warehouses that are often created by individual departments or divisions to aid in decision-making. Some products and services, such customer management, marketing, and finance, may necessitate the development of a data mart.

Metadata

In order to understand and locate data, users need to be familiar with the data warehouse system and its contents. Metadata, sometimes known as data about data, includes information such as data format, encoding/decoding techniques, data definitions, and domain limits. Included are topics relating to business, such as definitions, data quality alerts, organisational changes, rules and assumptions, and more [10]. In order to better understand what data is available, how to get it, what it signifies, when to use it, and what it does, metadata is a great asset to the business user. Data warehouses may be better understood with the help of metadata browsers.

Accessing and Analysing Data

Users are able to access the data through the front end of the business intelligence system. It's a collection of methods and tools that hide the technical aspects of data retrieval so that business users may access the data directly, interactively, or in batches. An someone without technical training should have no trouble navigating the interface because of the way it provides information in a straightforward, businesslike style [12]. To do this, we make use of business intelligence (BI) tools, which are a collection of programs that deliver a GUI along with a number of reporting and analytical capabilities for businesses [2].

Hardware and Software Assets

Resources for Hardware

Graphics processing units (GPUs) made by NVIDIA or Intel are popular among academic institutions, data scientists, and researchers parts used to speed up, test, and train deep learning algorithms. Graphics processing units (GPUs) speed up the completion of parallel computing processes for data, photos, movies, and graphics. Upcoming Nvidia AI Chip A little computer with a lot of computing power is the Jetson Xavier from Nvidia. It includes graphic and video processing units, an ARM 64 CPU with 8 cores, two NVDLA accelerators, and a Volta tensor core GPU. With its nine billion transistors and 30 watts of power consumption, it can do 30 trillion operations per second (TOPS). With its efficient core hardware components, Intel's AI Chip-Neural network processor offers versatile support for all deep learning primitives. It uses less electricity while scaling many computing nodes and optimising computation efficiency [3–5].

Software Resources

Tools and software for artificial intelligence and deep learning are detailed. A GPU and CNN library are part of the versatile and extendable Pylearn2 package for machine learning techniques. An open-source machine learning library is made available by Torch, an AI and DL tool. The scripting language that is utilised is Lua. Because of its adaptability and computing efficiency, it is one of the most used tools. The lack of support for reimplementing logic is a drawback, though. Technet is a brand-new, user- and admin-friendly open-source framework. You may swiftly evaluate mathematical expressions with the help of the Theano Python module. It involves multidimensional arrays. Theano It outperforms the CPU in data-intensive tasks because to its use of mathematical expressions. In order to train on massive datasets, deep learning networks are employed. Caffe is a free and open-source AI deployment framework that works with many different libraries, including MATLAB, CUDA, Python, C++, and many more.

This quick neural network is cuda-convent compatible, meaning it works with Python, C++, and CUDA. To do this, we employ convolutional neural networks. Among the many supported libraries of the open-source Deeplearning4j framework are those for C, C++, Java, and Scala. The distributed framework library may now take advantage of graphics processing units. Nearly every type of deep neural network is available. One open-source framework for numerical computing is TensorFlow. When a data flow graph provides a suitable representation for the calculation, it is employed. Python makes it run more quickly. Embedded systems are an option, and it provides C++ APIs and CUDA support [8].

Analysing Market Trends

The worldwide artificial intelligence market is anticipated to see substantial growth over the next five years as a result of its extensive use in several industries, including automotive, banking, healthcare, consumer electronics, and more. Market leaders in artificial intelligence (AI) like Google, IBM, Microsoft, and others are acquiring more and more AI startups, which is driving investment in the industry. An analysis of the worldwide AI market from 2012 to 2022 predicts a compound annual growth rate (CAGR) of more than 60% from 2017 to 2022. Analyses of worldwide AI sales show that natural language processing, video and audio processing, picture identification, and gesture control account for a disproportionately big portion of the market [9]. When it came to artificial intelligence in 2016, picture recognition was king. Presumably, it will uphold the following five years [2–3].

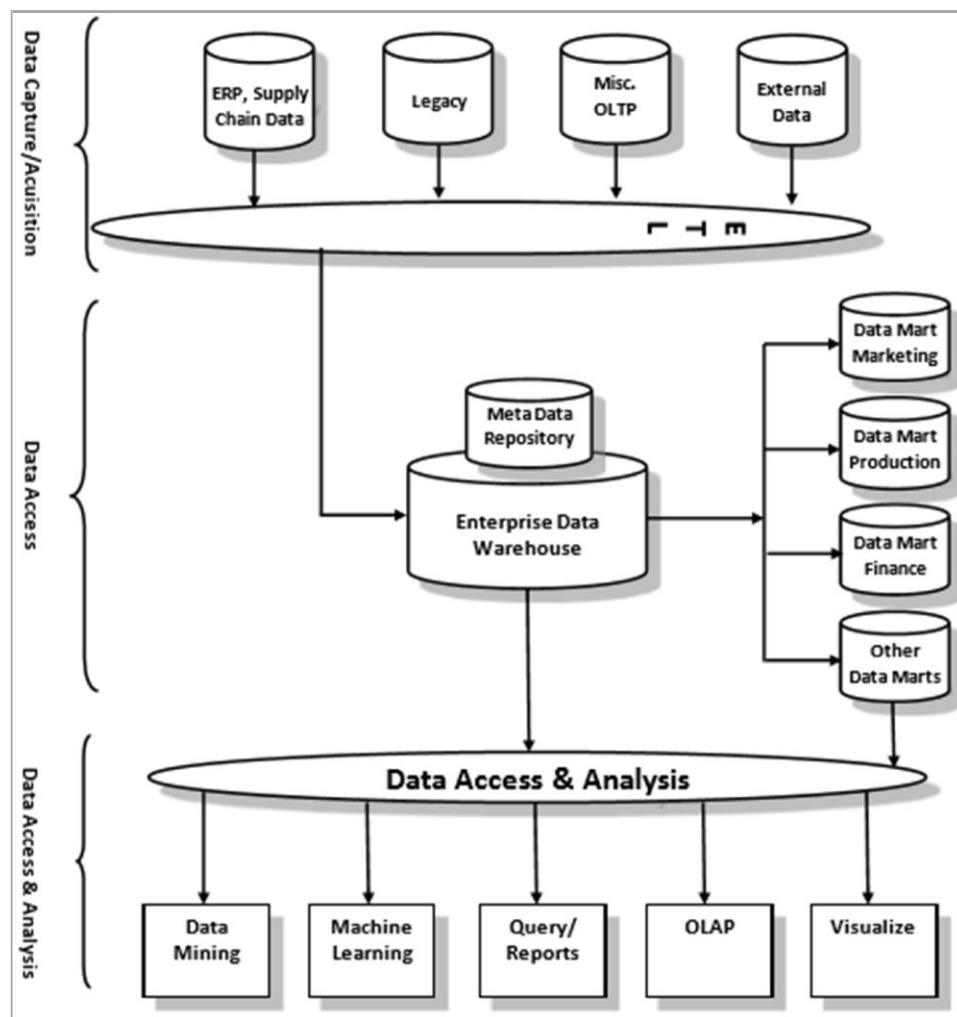


Figure 2. Framework of Business Intelligence (Valter et al., 2017).

Data Analytics and Data Science Research Challenges

Inadequate algorithms were utilised to solve big reasoning issues in the late 1980s and early 1990s because to the sheer number of conceivable problem combinations [3]. This resulted in an exponential decline in computing performance as the issues got bigger. Therefore, in order to cope with insufficient or unclear data, researchers in artificial intelligence came up with the ideas of probability and economics [4].

Safety

When presented with a dynamic environment or when physical touch is required for movement, object recognition becomes more challenging to program. The selection of effective solutions and the ability to have a substantial impact when integrating business intelligence (BI) in one's company can be challenging for executives [5].

Conclusion

The broad use of artificial intelligence is bringing us closer to its implementation in modern civilisation. It's equally important to acknowledge that these technologies may safely reduce dangers. Machine learning approaches are being used more and more for analytics forecasting. One argument in favour of deep learning over more conventional machine learning methods is the superior accuracy of its output. But the more progress we achieve in this area, the more challenging and intricate it gets. In today's cutthroat economic climate, the timely and accurate reporting of a company's financial data can spell the difference between success and failure. Business intelligence is invaluable, but no company can dispute that. Recent predictions from industry experts indicate that BI visual tools and analytics will be used every day by millions of people in the next years. Conclusions drawn from this research point to big data, BI, and AI as real foundational technologies for company growth in areas such as decision-making, forecasting, and the corporate economy. Additionally, the potential for development is greater when these technologies are supported by effective leadership and training.

References

- [1] Pasham, S.D. (2017) AI-Driven Cloud Cost Optimization for Small and Medium Enterprises (SMEs). *The Computertech*. 1-24.
- [2] Alam, K., Mostakim, M. A., & Khan, M. S. I. (2017). Design and Optimization of MicroSolar Grid for Off-Grid Rural Communities. *Distributed Learning and Broad Applications in Scientific Research*, 3.
- [3] Agarwal, A. V., & Kumar, S. (2017, November). Unsupervised data responsive based monitoring of fields. In *2017 International Conference on Inventive Computing and Informatics (ICICI)* (pp. 184-188). IEEE.
- [4] Mishra, M. (2017). Reliability-based Life Cycle Management of Corroding Pipelines via Optimization under Uncertainty (Doctoral dissertation).
- [5] Agarwal, A. V., & Kumar, S. (2017, October). Intelligent multi-level mechanism of secure data handling of vehicular information for post-accident protocols. In *2017 2nd International Conference on Communication and Electronics Systems (ICCES)* (pp. 902-906). IEEE.
- [6] Malhotra, I., Gopinath, S., Janga, K. C., Greenberg, S., Sharma, S. K., & Tarkovsky, R. (2014). Unpredictable nature of tolvaptan in treatment of hypervolemic hyponatremia: case review on role of vaptans. *Case reports in endocrinology*, 2014(1), 807054.
- [7] Shakibaie-M, B. (2013). Comparison of the effectiveness of two different bone substitute materials for socket preservation after tooth extraction: a controlled clinical study. *International Journal of Periodontics & Restorative Dentistry*, 33(2).
- [8] Gopinath, S., Janga, K. C., Greenberg, S., & Sharma, S. K. (2013). Tolvaptan in the treatment of acute hyponatremia associated with acute kidney injury. *Case reports in nephrology*, 2013(1), 801575.
- [9] Shilpa, Lalitha, Prakash, A., & Rao, S. (2009). BFHI in a tertiary care hospital: Does being Baby friendly affect lactation success?. *The Indian Journal of Pediatrics*, 76, 655-657.
- [10] Singh, V. K., Mishra, A., Gupta, K. K., Misra, R., & Patel, M. L. (2015). Reduction of microalbuminuria in type-2 diabetes mellitus with angiotensin-converting enzyme inhibitor alone and with cilnidipine. *Indian Journal of Nephrology*, 25(6), 334-339.
- [11] Gopinath, S., Giambarberi, L., Patil, S., & Chamberlain, R. S. (2016). Characteristics and survival of patients with eccrine carcinoma: a cohort study. *Journal of the American Academy of Dermatology*, 75(1), 215-217.