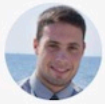


Initial Post

by Panagiotis Koilakos - Monday, 24 October 2022, 8:38 PM

Data volume production has increased exponentially throughout the years. While only two zettabytes of data were produced in 2010, in 2016, the data production volume reached 26 zettabytes, and by 2025 it is estimated to reach 181 zettabytes. With nearly 80 zettabytes in data connections projected by 2025 through IoTs, almost 50% of worldwide data volume will come from such systems (Statista, 2022).

Such volume presents unlimited opportunities but simultaneously creates many risks and challenges closely associated with technical limitations. As we live in an era where data protection is becoming more and more trending, with Data Protection Officers being required for every company under the General Data Protection Regulations applied in the EU and being embedded in the national legislation of most, if not all, EU countries. Ensuring that data are always safe, especially for interconnected and "always-available" devices, is one of the most critical data protection-related challenges (Ramachandran et al., 2018). In addition to the security issues, several constraints, such as bandwidth limits, processing speed, battery capacity and availability, especially in devices with integrated computing power, make using IoTs a decision that needs to be well thought out (Ramachandran et al., 2018; Dian et al., 2020).

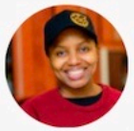
On the other hand, the same elements that create the challenges make using IoTs a popular option. Even though data protection presents a current challenge (coupled with the limited relevant regulation), anonymity can be achieved through blockchain technologies. Moreover, the broad spectrum of applications makes IoTs a go-to solution for many industries that aim for automation. Finally, the data production, linked with the relatively low cost of such devices, makes their use in many applications appealing, as the data wealth can be enormous, especially for machine learning applications (Ramachandran et al., 2018; Statista, 2022; Shafique et al., 2018).

Data show that IoT devices are used daily, assisting companies and individuals in their daily routines. Even though the opportunities are virtually endless, one must remember that such applications are not a panacea, and the advantages must be cogitated against the disadvantages.

Re: Initial Post

by Constantinos Kyriacou - Monday, 7 November 2022, 5:53 PM

Hello Panagioti,
As correctly stated within your post data protection is one of the most crucial challenges we have to keep in mind especially when dealing with big data.Therefore the strict application of the EU General Data Protection regulation is of the utmost importance. At the same time however, it is referred that some restrictions as such bandwidth limits , the rate at which the data is processed, and battery capacity and availability , come to add to the difficulty of processing large amount of data making it even more challenging. Concerning these restrictions ,are there any practical ways which can contribute in eliminating the difficulties we face when analysing this kind of data ?



Initial Post

by Tsitsi Mkonto - Sunday, 9 October 2022, 7:44 AM

The Internet of Things (IoT) has transformed the way organisations collect large amounts of varied data, store it and use it to gain analytical insights that give them a competitive advantage in the area of specialisation. However, for such insights to be availed and taken advantage of, data needs to be accessed timely, accurately, and efficiently. Data that is collected from IoT devices is usually incomplete, in a format that cannot be analysed or large segments of the data may be missing. In order to get the data to the analysis stage, data wrangling needs to be performed. When using IoT devices the process of data wrangling becomes more complex, as the number of connected devices grows daily and the amount of data sent by these devices grows exponentially by the hour. It's important to note the sheer amount of data collected and the types of data received in real-time, require tools that can process the data automatically without manual efforts (Selfridge,2016). The first challenge that organisations need to be cognisant of, is that the preparation of data should be rapid, thus reducing the time spent on data wrangling. The tool chosen to do the data wrangling should be federative, meaning it promotes and simplifies collaboration between analysts and data scientists (Schwab, 2021). This can be achieved by developing in house expertise in data wrangling who will be responsible for the task and building the collaborative systems.

For organisations that require real-time results, the lamda architecture approach solves this issue. The lambda architecture has several benefits such as flexible scaling, there is no need for server management and it allows for organisations to react in real-time to changing environments. However, it is a highly complex and administrators have to maintain two separate code bases for batch and streaming layers.

Re: Initial Post

by Constantinos Kyriacou - Monday, 7 November 2022, 3:58 PM

Hello Tsitsi ,
Processing large amount of data can be really challenging as it is mentioned within your post. In order to reach to the point where we can effectively analyse our data and make robust conclusions we come up against many challenges and risks.As rightly mentioned in your post ,data collected from IoT devices is usually incomplete, missing and complex due to the large size of the data.To deal with these kinds of problems it is required to use correctly some fundamental cleaning techniques.Claiming to the fullest the cleaning methods,gives us the ability to remove irrelevant ,bad data and duplicates and thus simplify our analysing process. In addition to that ,handling missing values contributes to fill out and complete our dataset in order to make it more suitable for our analysis. Additionally, as stated within your post,when we have to deal with big data we have to reduce as much as possible the time spent in preprocessing. Are there any handy ways which can contribute to accomplish our aim?