Machine learning has been around for several decades, but it has recently exploded due to much more data being available, more powerful computing, and more tools that are easier to use. Machine learning is different from traditional programming in that it allows a data scientist to create a learning algorithm that identifies patterns and uses these for future predictions.

For this task, you will research the application of machine learning to streaming data.

An example in your daily life is an online gaming company that collects streaming data about player–game interactions and feeds the data into its gaming platform. It then analyzes the data in real-time and offers incentives and dynamic experiences to engage its players.

* Describe the differences between streaming data and at-rest data.
* What are other streaming data examples and potential algorithms for use in your daily life?
* Would you recommend stream or batch processing for processing the big data collected from online gaming? Explain your decision.

**Differences between streaming data analytics and at rest data analytics**

* Streaming data is constantly arriving, while data at rest has a defined start and end period
* There is a much higher volume of streaming data than data at rest
* There is a low latency with streaming data (it arrives quickly and is analyzed quickly), while there can be a much slower analysis process with data at rest
* Streaming data can allow for real time actions to be taken based on business intelligence gained from analyzing it (movie recommendations, fixing machinery, etc.) while data at rest shows business intelligence after an event has happened.

(These bullet points are from my IP1 assignment for this class)

**Streaming Data Examples in my daily life**

* **Movie streaming services such as Peacock and Netflix:** recommendation algorithms as well as the streaming shows and movies themselves.
* **Oura ring activity mode:** A personal health tracker for sleep, activity, and wellbeing. It has a setting to track an activity where it continuously monitors me while walking, running or biking and gives me an overview of the activity: how far, how long, max speed, max heart rate, average heart rate, heart rate zones, a graph of your heart rate and your route (*Activity Heart Rate*, n.d.).
* **Metered grid connection:** A smart grid connection constantly monitors electricity usage for a household to determine peak usage periods, average usage, and predict the monthly bill.

**Stream vs batch processing for online gaming big data:**

Online gaming is not something I do very much (well, at all). Forgive any generalizations or mistakes I make when it comes to how online gaming works. From Weber (2018) it seems like using data streams as the sources of a data lake is a functional architecture for online gaming. You can set up topics for the game and events that occur in the game are captured, sent to the relevant topic, and then sent into a data lake in a semi-structured table format. From the lake, data can be queried to analyze the data. I assume some of these queries run on a continuous basis (streaming analytics) for real time feedback while playing the game, such as areas of the map you can access and powers/weapons you are eligible to use. Other of these queries could be used for after-the-fact analysis to answer questions such as: what pathways are users following to win the game? Which sections are too hard/too easy? Which characters are the most/least popular?

I might worry about processing delays in streaming analytics for online games. Sending so much metadata takes bandwidth, and with online games everything is in real time and it is important to not have lags or pixelated images. Making sure your pipelines are efficient and small enough to not affect the players is imperative. In conclusion, both real-time and batch analytics are useful in online gaming.

**References**

*Activity Heart Rate*. (n.d.). Oura Help. Retrieved October 25, 2024, from https://support.ouraring.com/hc/en-us/articles/6321008135699-Activity-Heart-Rate

Weber, B. (2018, April 1). *A Fully-Managed Game Analytics Pipeline*. Game Developer. https://www.gamedeveloper.com/programming/a-fully-managed-game-analytics-pipeline-