Group Members:
Dan Le, Kaitlyn Tom, Madison Long

Topic: QoE Prediction For Streaming

- **Motivation**

  Video streaming services such as Netflix, YouTube, and Twitch have grown exponentially over the past decade. User experience has become the main metric used to evaluate the performance of these streaming services. Traditional metrics like Quality of Service (QoS) give insight into system performance but do not accurately evaluate user satisfaction. Predicting Quality of Experience (QoE) is important for improving adaptive stream algorithms, optimizing network resource allocation, and identifying when users have poor streaming experiences. QoE prediction would bridge the gaps between network metrics and human perception, allowing for more user-centric streaming services.

- **Specifications**

  The goal of this project is to predict a user's QoE during a video streaming session using network metrics, session history, and playback context as input. Given a set of parameters such as throughput, packet loss rate, and buffer occupancy, the model will predict the outcome of QoE, which can either be a score from regression or a class from classification. The main objective is to reduce the error between predicted and actual QoE.

- **Existing Approaches**

  Some existing approaches include the usage of Machine Learning models using Random Forests for prediction of the QoE [1]. They also use chunk-based traffic analysis to extract lightweight features from encrypted flows, enabling real-time inference of video playback metrics such as buffering, video state, and resolution. Tools like NetMicroscope [2] already predict QoE from basic network features, and models like Pensieve [3] use reinforcement learning to optimize streaming QoE by controlling bitrate. Specifically, Pensieve trains a neural network to select bitrates based on real-time network and buffer observations [3]. However, many existing approaches focus on specific platforms or settings and do not always offer generalizable or explainable models.

- **Existing Datasets**

  Some existing datasets that we can work on are from the Puffer project (Stanford's open streaming platform) [4] and netMicroscope (network telemetry datasets for QoE prediction) [2]. Additionally, by using netReplica [5], we can collect or simulate new datasets under varied network conditions to expand our training data.

- **Novelty**

  As we have discussed in class, a common limitation for many network machine learning tends to be a lack of replicability and generalization that occurs once the model leaves the lab specific conditions on which it was trained. Our project intends to address this issue by including controlled datasets, such as those from Stanford's Puffer project and Requet) as well as various data points extracted from our own experiments using netReplica. These will be used to train a more robust and generalizable QoE prediction model. Theoretically breaking away from a single static laboratory environment. Additionally, we will integrate model explainability techniques to increase transparency potentially utilizing SHAP or TRUSTEE to be determined based on our experience.

- **Implementation**

  Our approach will use netMicroscope to monitor and collect key metrics that reflect users quality of experience during streaming sessions of the video streaming platform Puffer. In addition to Puffer data, we will collect network data from multiple sources, including our own measurements via Requet and simulated sessions via netRepilca. Using this dataset we will train a machine learning model to predict QoE outcomes such as rebuffering events, playback interruptions, or resolution drops. Metrics for success can be measured via F1 score for classification, mean squared error for regression and QoE specific metrics such as correct prediction of rebuffering events. Ideally we would like to maximize scores in all of these fields and the project will be considered successful if the model is able to correctly predict the users quality of experience based on the provided condition in real time simulations via netReplica. Aka, achieve high performance in various field conditions, as opposed to limited lab based validation results. This model will allow us to interpret these results and identify the network features most relevant to understanding prediction behaviors.

References:

[1] Craig Gutterman, Katherine Guo, Sarthak Arora, Xiaoyang Wang, Les Wu, Ethan Katz-Bassett, and Gil Zussman. 2019. Requet: real-time QoE detection for encrypted YouTube traffic. In Proceedings of the 10th ACM Multimedia Systems Conference (MMSys '19). Association for Computing Machinery, New York, NY, USA, 48–59. https://doi.org/10.1145/3304109.3306226

[2] NetMicroscope Team. *NetMicroscope: Real-Time QoE Detection from Encrypted Traffic*. [Online]. Available: https://www.netmicroscope.com/.

[3] Hongzi Mao, Ravi Netravali, and Mohammad Alizadeh. 2017. Neural Adaptive Video Streaming with Pensieve. In Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM '17). Association for Computing Machinery, New York, NY, USA, 197–210. https://doi.org/10.1145/3098822.3098843

[4] Francis Y. Yan, Hudson Ayers (Stanford University), Chenzhi Zhu (Tsinghua University), Sadjad Fouladi, James Hong, Keyi Zhang, Philip Levis, and Keith Winstein (Stanford University). *Learning in situ: a randomized experiment in video streaming*. In *Proceedings of the 18th USENIX Symposium on Networked Systems Design and Implementation (NSDI '21)*, 705–720. USENIX Association, 2021. https://puffer.stanford.edu/static/puffer/documents/puffer-paper.pdf

[5] NetReplica. *Network Replica: Predicting Quality of Experience from Encrypted Traffic*. GitHub repository. Available at: https://github.com/netreplica.