Appendix A. Dataset Details

We use Avazu, Criteo-Kaggle and Criteo-Terabyte, three large popular open source benchmark CTR datasets to evaluate MEM-REC. Table 1 provides the details like the number of data points, alphabet size, number of dense and sparse features etc. of these datasets. Each record has numeric and categorical features describing an online advertisement served to users as a web/mobile advertisement. The goal is to predict whether or not a user clicks on an ad.

	Avazu	Criteo Kaggle	Criteo Terabyte
Data points (Millions)	40	46	430
Alphabet Size (Millions)	9	34	188
num_dense_features	0	13	13
num_sparse_features	22	26	26
num_days	10	7	24

Table 1: Dataset properties

Appendix B. Dataset Partitioning

The dataset contains n consecutive days of data, each with an approximately equal number of samples. Typically, data points from the first $(n-1)$ days are used as the training set while the data points from the nth day are split into a validation and test set. We follow this approach for partitioning the data as well.

Appendix C. Hyperparameters

We follow the basic setup like optimizer, sparse feature size, MLP architecture as suggested by the DLRM paper and their GitHub implementation. These parameters are summarized in Table 2.

hyperparameter	value	
	Criteo-Kaggle /Avazu	Criteo-TB
sparse-feature-size	16	128
arch-mlp-bot	13-512-256-64-16	13-512-256-128
arch-mlp-top	512-256-1	1024-1024-512-256-1
loss-function	bce	bce
optimizer	SGD	SGD

Table 2: DLRM Baseline Hyperparameters