AC Detection Validation

Weiqi Wang, Zixuan Zhou

2022/6/2

Unsupervised -> Supervised Learning

- Measured the efficiency of each air conditioner (Calculation at here)
- Calculated the mean of all recorded efficiency, as the borderline to classify normal AC and poor AC.
- Time period: 2020/12/30 – 2021/08/15

Agenda

- 1. Dataset Loading Technique
- 2. Methodology
- 3. Tasks and Settings
- 4. Result
- 5. Conclusion & Future work
- 6. TODO

1. Dataset Loading Technique – I

- Sparse Dataset: manually setting **number of data** and **group size**
- Generating one sparse data by sampling *group size* amount of data from 80% randomly sampled data per room. Other 20% are used to sample validation data. ***No duplicate between two sets!***
- The original data follows its chronological order in each data sample (from early to late)
- **Number of sparse data to be sampled from its room is linearly correlated to the room's total amount of data.**
- May not be fully generated, for example, for group size 200, we can only sample 32,450 data from all rooms for training, and 8,100 for validation.
- If we insist on sampling sufficient number of data, there will be repeating data from the same room, which is not ideally what we want in machine learning.

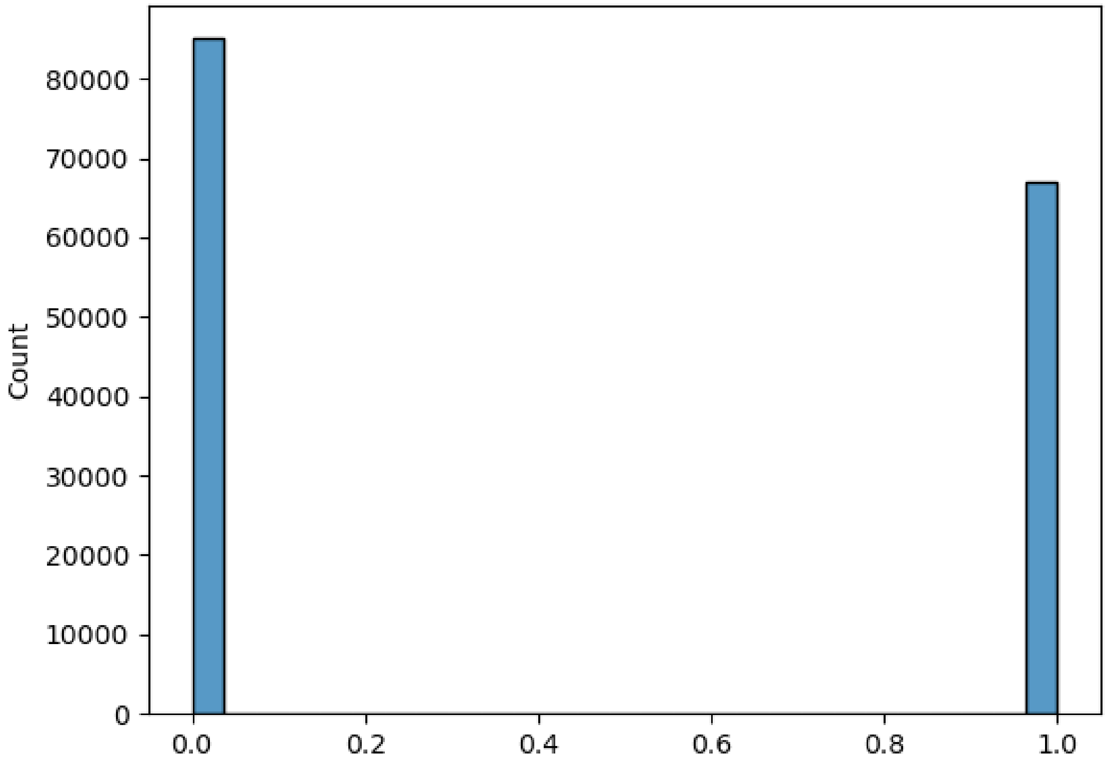
1. Dataset Loading Technique – II

- For small group sizes, such as 5, 10, we can expect a total amount of training data at 400,000.
- For middle sizes, such as 25, 50, decreases to 100,000 – 150,000
- For large sizes, such as 100, 200, 400, less than 50,000, even 10,000.

- Training: Validation=8:2, can also estimate validation data from above

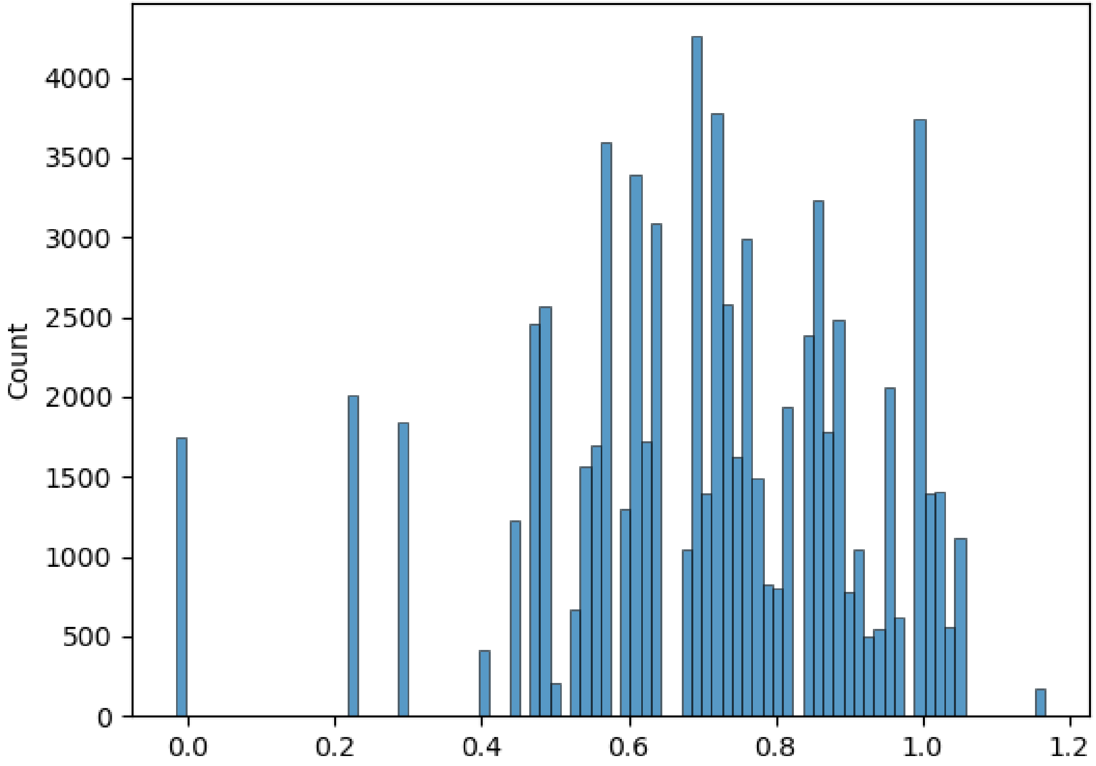
1. Dataset Loading Technique – III

- Classification data distribution: Group size 50 (training)



1. Dataset Loading Technique – IV

- Regression data distribution: Group size 50 (training)



2. Methodology – encoder selections

- 1. LSTM
- 2. Bi-LSTM
- 3. Transformer
- 4. LSTM-Transformer
- 5. BiLSTM-Transformer

- Unsupervised baseline: XGBoost

2. Methodology – model architecture

- Step 1: processed by encoder to generate embedding
- Step 2: four layers of fully-connected neural network:
  - Dense to a 512-dim vector, batch norm, dropout, activated ReLU
  - Dense to a 128-dim vector, batch norm, dropout, activated ReLU
  - Dense to a 64-dim vector, batch norm, dropout, activated ReLU
  - Dense to a 2-dim vector or a 1-dim scalar, depending on the task
- Step 3: SoftMax to generate a ***probability vector*** for classification

  ReLU to generate a ***positive*** prediction for regression

2. Methodology – training details

- All models trained for 200 epochs, with batch size 64
- AdamW optimizer, default learning rate $5 * 10^{-5}$
- Linear scheduler, decrease the learning rate linearly by each iteration (batch).
- Deep Learning Framework: PyTorch 1.10.0, with support of the huggingface library.

- Metrics for classification: accuracy, f_1 score.
- Metrics for regression: RMSE, R2 score

3. Tasks and Settings – two settings

- Setting 1: Generate sparse data from **all** rooms. For each room, 80% of data are used for training and other 20% for validation.
  - Advantage:
    - training data: validation data=4:1 is strictly guaranteed
    - performance is outstanding
    - Traditional setting, prove the plausibility of our method. Justifiable because training data has no duplication with validation data. Didn't peek the answers.
  - Disadvantage:
    - Need previous data for each room to achieve high performance.
    - Generalizability? How's the performance when encountering unseen rooms?

3. Tasks and Settings – two settings

- Setting 2 – **Zero-Shot Experiments**: Generate training sparse data from **80%** rooms. For each room, all of data are used. For the other **20%** rooms, use them all to generate validation data.
  - Advantage:
    - Can be used to prove generalizability. *More suitable for real world applications?*
    - Can release the model as pretrained version. Other organizations/researchers can directly download the model and run on their data.
  - Disadvantage:
    - Performance is only satisfiable
    - Data number is NOT guaranteed. For example, randomized sampling 80% rooms can sample all the rooms with large amount of data, validation data may be only a little.
    - But we cannot fix the rooms for training data to achieve certain ratio, violate the human intervention rule for machine learning data preparation.
    - What we can do: use small group size and large data number, so that even for rooms with few data, we can still sample as many data as possible.

3. Tasks and Settings – two tasks

- Task 1: AC Efficiency Classification
  - Given a sparse data x, predict whether the AC is normally efficient or inefficient.
  - Label is determined by whether the AC efficiency is larger or lower than mean efficiency of all ACs.

- Task 2: AC Efficiency Regression
  - Given a sparse data x, predict the efficiency of the AC.
  - Label is directly retreived from the calculated efficiency.

4. Results – Experiments

- The number of experiments is huge, there can be at most: $8 * 8 * 5 * 4 = 1260$ experiments. Can reduce to around 900. Currently finished about 100.
- https://github.com/MighTy-Weaver/AC-Detection-Validation

Data Number - Group Size	5	10	25	50	100	150	200	400
10,000				5 Done				
20,000								
50,000				5 Done				
100,000		5 Done	5 Done	5 Done	5 Done	5 Done	5 Done	
150,000								
200,000			5 Done	5 Done	5 Done			
300,000								
400,000			5 Done	5 Done	5 Done			

4. Results – Baseline

- XGBoost baseline, predict every half hour’s AC consumption. Then Shapley plot, then k-means.

- Classification task:
  - **Accuracy: 0.5824**
  - $\boldsymbol{f_1}$ **score: 0.5766**

- Regression task:
  - **RMSE: 0.1877**
  - **R2 score: 0.5109**

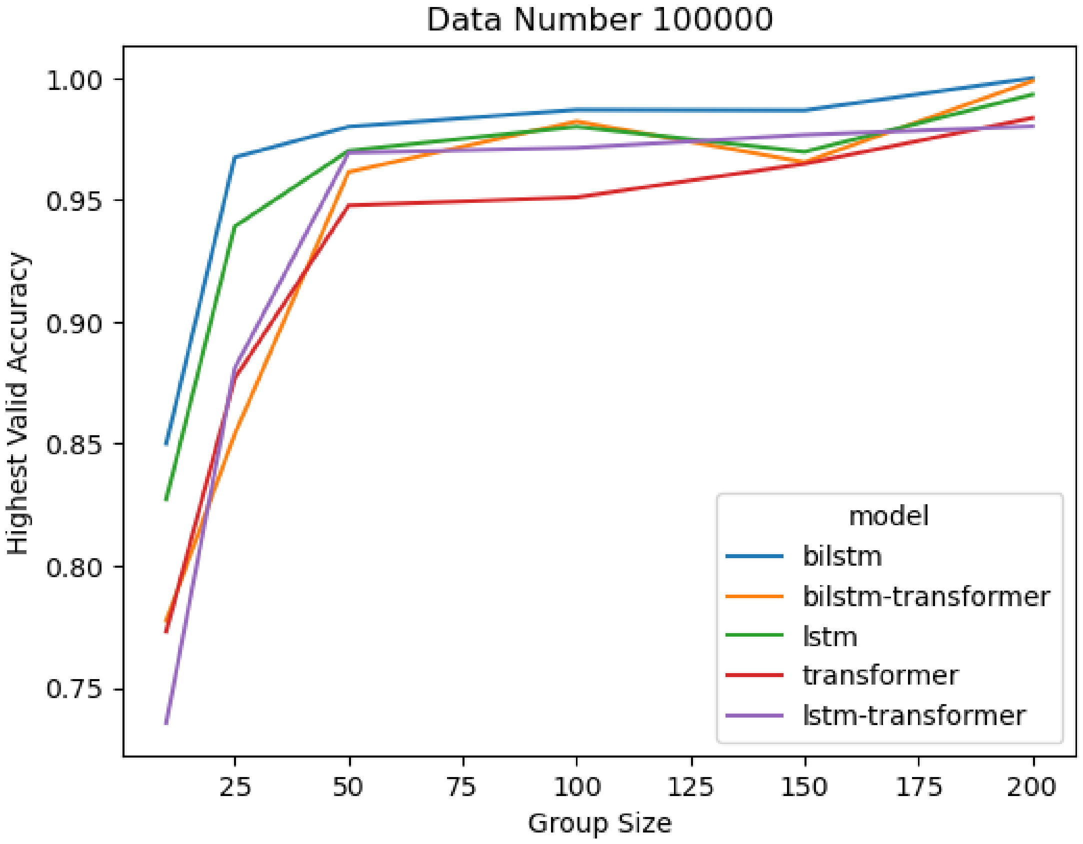
4. Results – Experiment results

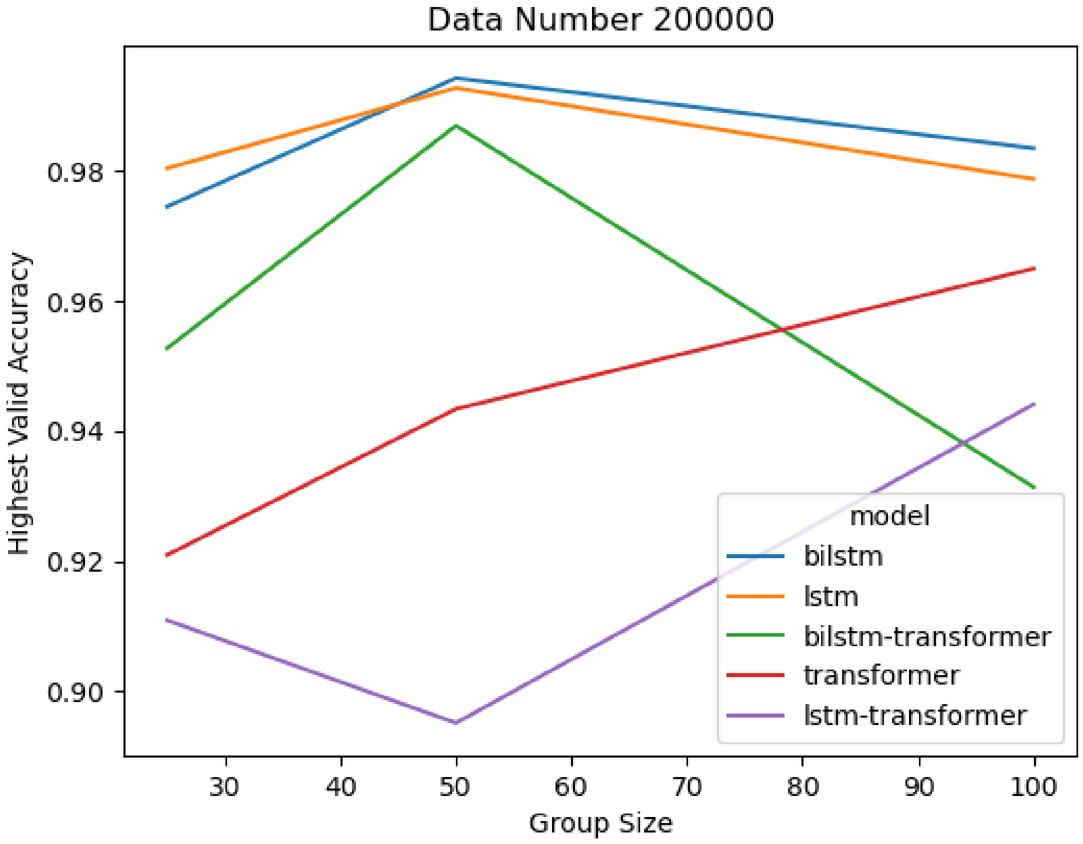
- Currently, only setting 1’s classification results can be directly viewed, as 70 experiments are finished.
- Can be downloaded at:
  - https://mighty-weaver.github.io/files/sparse_record.csv
- Can be directly viewed at:
  - https://github.com/MighTy-Weaver/AC-Detection-Validation/blob/main/sparse_classification/sparse_record.csv

model	gs	data_number	best_train_acc	best_valid_acc	best_train_f1	best_valid_f1
bilstm	200	100000	1.0	1.0	1.0	1.0
bilstm-transformer	200	100000	1.0	0.9987702902115101	1.0	0.9984040855410151
bilstm	50	400000	0.99943413979379	0.9977909851154473	0.9993572688409912	0.9974943324185658
bilstm	50	200000	0.9976056727138779	0.9941677175283733	0.9972804972804974	0.9933932504017617
lstm	200	100000	0.9999076184024143	0.9932365961633055	0.999879995199808	0.9912712267893985
lstm	50	200000	0.9948298317393077	0.9926439680538042	0.9941275794569878	0.9916755856820074
lstm	50	400000	0.9978220961830755	0.9924262346815337	0.9975263618088198	0.9914163090128755
bilstm	25	400000	0.9900580868499456	0.9874156483676819	0.9887847030704173	0.985823721053095
bilstm	100	100000	0.9995112149025145	0.9869946894982118	0.9994426899498422	0.9850783387217109
bilstm-transformer	50	200000	0.9996711088892689	0.9868642286675073	0.9996264921638056	0.9851843072182055
bilstm	150	100000	0.9997756086615057	0.9867861142217246	0.9997365474543898	0.9847150259067357
bilstm-transformer	25	400000	0.9983506203101877	0.9864516297126182	0.9981396922036192	0.9848343443770416
lstm	25	400000	0.9790470105808033	0.9855657747322895	0.976368342107585	0.9837543541258087
transformer	200	100000	1.0	0.9836448598130841	1.0	0.978912319644839
bilstm	100	200000	0.9933441999456669	0.9833957349829073	0.9924118066094713	0.9812339016313013
bilstm-transformer	100	100000	0.9999185358170858	0.982226075647556	0.9999071236184639	0.9797580844235991
lstm	25	200000	0.9704379562043796	0.9803247970018738	0.9667019027484144	0.9777070063694266
lstm-transformer	200	100000	1.0	0.9802016724053123	1.0	0.9747609343157234
bilstm	50	100000	0.9933203597538526	0.9800671422576585	0.9924197206637221	0.9778451492537313
lstm	100	100000	0.9969586705045348	0.980058523897258	0.9965337954939342	0.9775555013417907
lstm-transformer	50	400000	0.9986708864923904	0.9789354652080156	0.9984905096398148	0.9764736981231826
lstm	100	200000	0.993493615865254	0.9786749145368713	0.9925827281314359	0.9757960214325307
lstm-transformer	150	100000	0.9997756086615057	0.9767077267637178	0.9997365301014359	0.9729025534132361
bilstm-transformer	50	400000	0.999888143912726	0.9750433913638037	0.9998729551382173	0.9719115963722815

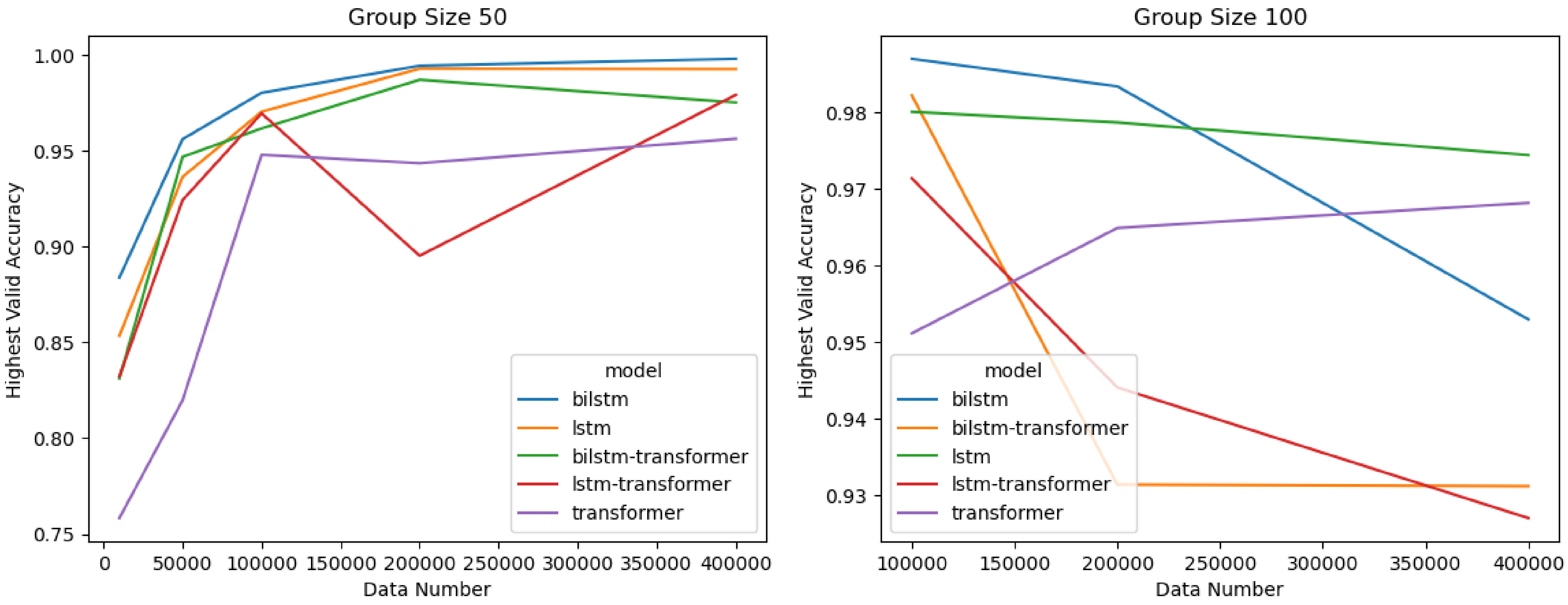
44	bilstm-transformer	50	50000	0.9726250525430853	0.9466861802216182	0.9689085158441249	0.9404622927854309
45	lstm-transformer	100	200000	0.9819206737299647	0.9440555645992729	0.9793978794210975	0.9365733620424483
46	transformer	50	200000	0.9921592359201715	0.9433585540142917	0.9910951740624533	0.9375
47	lstm-transformer	25	400000	0.988584727917544	0.9416117349730335	0.9871324745184783	0.9364813922507865
48	lstm	25	100000	0.9140460123498606	0.9391105569409809	0.9027732751760927	0.9315260574900678
49	lstm	50	50000	0.9399957965531736	0.9362324900689943	0.9320480780673569	0.9305713635328933
50	bilstm-transformer	100	200000	0.9856424884542244	0.931358185468555	0.9836451128750252	0.9161974163630341
51	bilstm-transformer	100	400000	0.9769228678958166	0.9311627654374627	0.9737080408349651	0.9205453690644101
52	lstm-transformer	100	400000	0.9548987106153449	0.926980937381198	0.9485759432391192	0.9193678970883686
53	lstm-transformer	50	50000	0.962589323245061	0.9241062094919507	0.9575533563848813	0.9150479756611279
54	transformer	25	200000	0.9304874869655891	0.9209348323964189	0.9217656637375857	0.9089055472263868
55	lstm-transformer	25	200000	0.9507429614181439	0.9108890276910264	0.9444354589698725	0.9036903690369037
56	transformer	25	400000	0.9748029545410096	0.9050050806388578	0.9716119839293715	0.888046346567988
57	lstm-transformer	50	200000	0.996961046136845	0.8951240016813787	0.9965480655718107	0.8841017303449077
58	bilstm	50	10000	0.9134966128191766	0.8837209302325582	0.9017751479289942	0.8704318936877077
59	lstm-transformer	25	100000	0.901826424532972	0.881026600166251	0.8885998107852413	0.8620162333843077
60	transformer	25	100000	0.8987780412183112	0.8767664172901081	0.8857184880129431	0.8523038605230385
61	bilstm-transformer	25	100000	0.9192829785570986	0.8539068994181214	0.9089894242068155	0.8398268398268398
62	lstm	50	10000	0.8402813965607087	0.8533872598584429	0.8199706314243759	0.8394241417497232
63	bilstm	10	100000	0.8371960239806908	0.8500620860927153	0.8135643593782507	0.8270676691729323
64	lstm-transformer	50	10000	0.8421052631578947	0.8321536905965622	0.8172237277928335	0.8113636363636363
65	bilstm-transformer	50	10000	0.8540906722251173	0.8311425682507584	0.8306110102843316	0.8265835929387332
66	lstm	10	100000	0.8007578313565701	0.8272971854304636	0.7691246588898553	0.7947260191121326
67	transformer	50	50000	0.937526271542665	0.8197783817687644	0.928985247566147	0.7852882703777336
68	bilstm-transformer	10	100000	0.7626326853702213	0.7775248344370861	0.7156271376158199	0.7271351078592153
69	transformer	10	100000	0.7496042147880927	0.7730753311258278	0.6931297709923664	0.743162108072048
70	transformer	50	10000	0.9616988014590933	0.7583417593528817	0.9568281938325991	0.7333333333333334
71	lstm-transformer	10	100000	0.7369131348784096	0.7355132450331126	0.6741450244278735	0.7313903111461205

4. Results – Data Number Fixed





4. Results – Group Size Fixed



4. Results – Setting 1 Regression Task

- Currently, only tested five models' performances at data number of 100,000 and a group size of 50.

- Best model (Bi-LSTM)'s performance:
- Max validation R^2 score: 0.8914
- Min validation RMSE: 0.07239

4. Results – Setting 1 Comparison

- Classification task: (in terms of the best validation performance)

	Accuracy	F_1 score
XGBoost	0.5824	0.5766
Sparse Deep Learning	**1.0**	**1.0**

- Regression task: (in terms of the best validation performance)

	RMSE	R^2 score
XGBoost	0.1877	0.5109
Sparse Deep Learning	**0.0724**	**0.8914**

4. Results – Setting 2 Result

- Performance is far below setting 1.
- Due to limited time and computational power, only tested 5 models' performances on 2 tasks at the data number of 100,000 and group size of 50.
- Classification task:
  - Max Validation Accuracy: 0.7200
  - Max Validation F_1: 0.6274
- Regression task:
  - Max validation R^2 score: 0.3994
  - Min validation RMSE: 0.2619

4. Results – Setting 2 Result Comparison

- Compared with Setting 1's results. Classification is acceptable, but regression is poor.

	Accuracy	F_1 score
XGBoost	0.5824	0.5766
Sparse Deep Learning Setting I	**1.0**	**1.0**
Sparse Deep Learning Setting II	0.7200	0.6274

	RMSE	R^2 score
XGBoost	0.1877	0.5109
Sparse Deep Learning Setting I	**0.0724**	**0.8914**
Sparse Deep Learning Setting II	0.2619	0.3994

5. Conclusion – Summary

- Supervised approach is significantly better than unsupervised learning.
- Sparse data sampling is extremely effective. It’s the key to setting 1’s success. In the traditional machine learning setting, our model can conquer both classification and regression tasks.
- From zero-shot experiments, our model has a certain degree of generalizability, though its performance is not dominative. However, it’s still a signal of our model is trying to capture the pattern of climate & electricity data, rather than remembering them.

5. Conclusion – Application

- For organizations/researchers with all required data and some ACs' measured efficiency, can finetune our model and effectively detect unknown efficiency ACs' efficiency level, or predict AC's efficiency.

- For organizations/researchers who don't have any efficiency available. We can release our pretrained model for them to directly make predictions. However, the performance is not guaranteed, as data from different areas are different. The model is more likely to be able to detect AC from humid or hot areas (such as Hong Kong)

5. Conclusion – Future work

- Generalizability is still not optimized; future researches can follow up.

- If other researchers have larger scale dataset, such as indoor & outdoor data for a certain period, can use our methods to train a dominating pretrained model that can generate very representative AC embeddings.

6. TODO

- Complete the remaining experiments (may need a month).
  - Setting 1’s classification & regression
  - Setting 2’s classification & regression
  - The general methodology, proposed method, contributions are unchanged. Need more experiments results to support.

- Select a target journal and start drafting our paper.