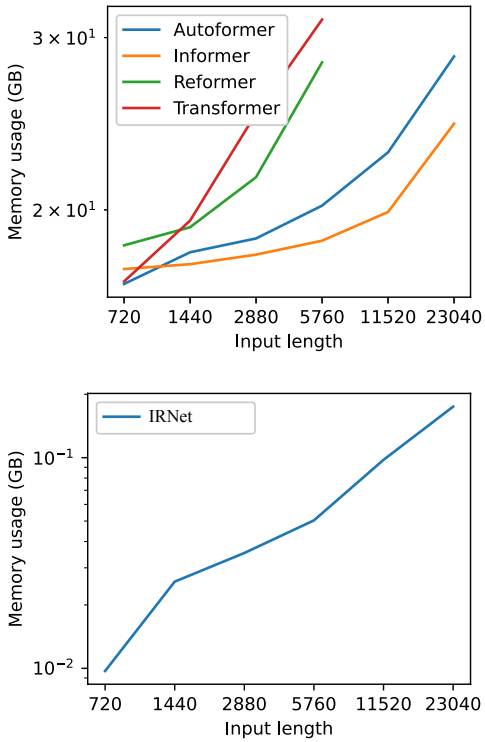

Figure 5: Left: the memory usage of Autoformer, Informer, Reformer and Transformer models with prolonging input length. Right: the memory usage of IRNet with prolonging input length.

three components is similar. It is worth noting that for long-range 720 predictions in *Exchange*, any two combination of components does not achieve excellent results, showing that the prediction results are better when the three components work together.

N SENSITIVITY ANALYSIS

In this part, we examine the performance sensitivity of the hidden layer dimension (shown in Table 12) and the depth of tree (shown in Table 13) in the IRNet model.

As shown in table 12, firstly, with the number of hidden dimension increasing, the prediction accuracy improves. After reaching a certain number (such as 128), the subsequent increase of hidden dimension would make model begin to overfit the training data. Therefore, the tradeoff between the complexity and performance of our model can be controlled to a suitable hidden size. The depth of our tree-based model has a similar impact as shown in table 13. The performance reaches a saturation status when the depth increases to some point, which shows that our model does not require a large depth to achieve more accurate predictions.

Table 12: Sensitivity experiments of hidden dimension. The best results are highlighted in bold.

Hidden Dim		32		64		128		256		512	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm2	96	**0.174**	0.264	0.192	0.292	0.179	**0.245**	0.175	0.266	0.199	0.291
	192	0.248	0.317	0.282	0.352	**0.233**	**0.309**	0.237	0.309	0.241	0.309
	336	0.308	0.364	0.319	0.377	**0.297**	**0.351**	0.306	0.364	0.286	0.341
	720	0.600	0.557	0.547	0.510	**0.377**	**0.395**	0.406	0.432	0.432	0.450
Exchange	96	0.098	0.226	0.108	0.235	0.091	0.224	**0.087**	**0.212**	0.089	0.216
	192	0.248	0.371	0.244	0.367	**0.208**	**0.335**	0.246	0.380	0.227	0.350
	336	0.448	0.514	0.348	0.441	0.345	0.440	0.408	0.484	0.447	0.510
	720	1.456	0.862	1.215	0.844	**0.690**	**0.647**	1.848	0.971	1.857	0.973

Table 13: Sensitivity experiments of different tree depths. The best results are highlighted in bold.

Tree depth		1		2		3	
Metric		MSE	MAE	MSE	MAE	MSE	MAE
ETTm2	96	0.203	0.283	**0.179**	**0.245**	0.189	0.285
	192	0.246	0.317	**0.233**	**0.309**	0.279	0.347
	336	0.320	0.374	**0.297**	**0.351**	0.314	0.364
	720	0.445	0.473	**0.377**	**0.395**	0.393	0.424
Exchange	96	0.136	0.270	0.091	0.224	**0.089**	**0.218**
	192	0.224	0.349	0.208	0.335	**0.198**	**0.327**
	336	0.458	0.524	**0.345**	**0.440**	0.506	0.531
	720	0.768	0.721	**0.690**	**0.647**	1.673	0.916

O COMPARISON OF MEMORY USAGE

In this part, we examine the memory usage of different SOTA Transformer based models and our IRNet as shown in Fig. 5. It can be seen that the memory usage of our IRNet model is two orders of magnitude smaller than various transformer counterparts. Furthermore, the memory usage of our IRNet increases **linearly** with prolonging input length, compared to $O(nlogn)$ increasing rate of the Transformer based models.