## Report of the Phase 1 of the Machine Learning Project (Quantum Thermodynamics)

Pooya Hosseini Kavyan Zare

#### 1 Report

The data is produced through a simulation. We produced one hundred thousand random Hamiltonians and thermal states (with random temperatures). We aim to build a machine to be able to predict the temperature given a Hamiltonian and a thermal state relating to that Hamiltonian. In order to so, we need to save Hamiltonians and the corresponding density matrices. But since not all the elements of these matrices are independent, we only saved the free parameters of each. Now we aim to give the free parameters of the Hamiltonians and the density matrices to the machine, and get the temperature associated with that state. Then we can compare the obtained temperature via our NN and the temperature defined according to this paper[1].

## 2 Sample Data

 $\begin{array}{l} \operatorname{Hamiltonian\ and\ density\ matrix} = [\ 1.81526199\ 0.06378736\ -0.98290981\ 1.12344083\ -0.66113381\ 1.92556437\ 0.32964225\ -0.5556897\ -0.40895781\ 1.56425678\ -1.53448789\ -0.42237833\ -0.62420827\ -0.10092788\ 0.04704152\ 0.75708847\ 1. \ -0.02731434\ 0.10834203\ -0.11633975\ 0.06204935\ -0.19097145\ -0.01616185\ 0.05187094\ 0.05690463\ -0.16198718\ 0.15210365\ 0.03925648\ 0.07319763\ 0.01994311\ 0.00315615\ -0.05039598] \\ \operatorname{beta} = 0.4386435209429565 \end{array}$ 

### 3 Statistical Analysis

We provide the results here. The file containing the codes is provided in our github page.

#### 3.1 Distribution of Features

The distribution of the first free parameter of the Hamiltonian, as an example, is available in our github page.

The distributions of other features are more or less the same, as we expect from our domain knowledge and the way we produced our data.

#### 3.2 Variance of Features

The variances of the free parameters that are acquired using Pauli operators that are not diagonal in the basis, are almost equal to 1.2 while the other parameters' variances are equal to 1.34. For the density matrices, the variances of the free parameters that are acquired using Pauli operators that are not diagonal in the basis, are almost equal to 0.02 while the other parameters' variances are equal to 0.023.

# 3.3 Feature Importance Based on How Important They Are in Predicting Labels

We trained a random forest model using our data. We observed that the importance of the Hamiltonians parameters were less and that is peculiar but we know regardless of this analysis that all the parameters are necessary so they are all important. Below, the first 16 numbers are the feature importance of the free parameters of Hamiltonians, and the next 15 numbers are for density matrices.

 $\begin{array}{c} [3.96676186\mathrm{e}\hbox{-}03, 6.80059794\mathrm{e}\hbox{-}03, 9.53791971\mathrm{e}\hbox{-}04, 1.65418991\mathrm{e}\hbox{-}01, 3.39072182\mathrm{e}\hbox{-}03, 2.62553653\mathrm{e}\hbox{-}03, 5.36514171\mathrm{e}\hbox{-}04, 2.40906940\mathrm{e}\hbox{-}03, 1.66656327\mathrm{e}\hbox{-}03, 4.28830789\mathrm{e}\hbox{-}04, 9.13109771\mathrm{e}\hbox{-}04, 1.41652305\mathrm{e}\hbox{-}03, 2.84079810\mathrm{e}\hbox{-}01, 2.58101284\mathrm{e}\hbox{-}03, 9.68498781\mathrm{e}\hbox{-}04, 1.82748012\mathrm{e}\hbox{-}01, 7.59131133\mathrm{e}\hbox{-}02, 9.30233177\mathrm{e}\hbox{-}03, 7.38070975\mathrm{e}\hbox{-}01, 4.11449022\mathrm{e}\hbox{-}02, 1.58894974\mathrm{e}\hbox{-}02, 4.73298378\mathrm{e}\hbox{-}03, 3.21930806\mathrm{e}\hbox{-}02, 1.42506619\mathrm{e}\hbox{-}02, 6.35491005\mathrm{e}\hbox{-}03, 1.11518673\mathrm{e}\hbox{-}02, 1.36579893\mathrm{e}\hbox{-}02, 9.34233554\mathrm{e}\hbox{-}01, 3.33882329\mathrm{e}\hbox{-}02, 7.70333814\mathrm{e}\hbox{-}03, 7.69163047\mathrm{e}\hbox{-}01] \end{array}$ 

#### 3.4 Correlation Between Different Features

The correlation matrix is produced in the code. As can be seen, different features do not have a significant correlation with each other, which is expected since they are the independent parameters as we generated the data.

So, using this and also our domain knowledge we may not use feature transformation and PCA.

#### 4 Conclusion

All in all, as can be inferred from our statistical analysis and also our domain knowledge and specifically the way we produced our data and saved our features, we might not have to manipulate our data before the training. The main reason is that our matrices were randomly generated. Moreover, we have kind of done this "feature engineering" before, since our features are the independent

parameters of these matrices, which are expected to be independent from each other.

## 5 References

[1] S. Alipour, F. Benatti, M. Afsary, F. Bakhshinezhad, M. Ramezani, T. Ala-Nissila, and A. T. Rezakhani(2021). Temperature in Nonequilibrium Quantum Systems