proposal of project 10

- **Group Member: Minke Yu (**minke.yu@duke.edu**) , Ziling Yuan(**ziling.yuan@duke.edu**)**
- **Project Name:** Neural Network Distillation
- **Heielmier cataclysm**
  - **What are you trying to do?** 1）Implement the distillation and investigate several properties of it；2) Reverse the student model and the teacher model to verify if the distillation still works；3）Implement self-distillation to prove that model can also learn from itself；4）Compare the results of distillation with and without data augmentation.
  - **How is it done today?** Study the concepts and logics of Distillation, and the implementation of it.
  - **Your approach and why do you think it will be successful?** Use Resnet20 as teacher model on CIFAR10 and use Resnet11 as student model. Train Resnent20 to get a good teacher model. Train a Resnet11 to see its accuracy on cifar10. Train another Resnet11 with knowledge distillation from Resnet20 and see if the accuracy is better than the previous one. Explore the soft target loss functions: KL-divergence, cosine similarity、MSE. Describe and explain all design choices.
  - **What are the risks?** 1) Student model is too small to convey the complex knowledge of teacher model; 2) There are two models (add a stage where teacher model is used to predict), which may lead to more training time and more storage space; 3) It is hard to control the ratio of soft targets and hard targets.
  - **What are the final "exams" to check for success?** Code implementing the method, table showing comparison results.