1) Choose and try to answer the "clarifying question" from one of the classmates in your group

Paper: Donti, Priya, Brandon Amos, and J. Zico Kolter. "Task-based end-to-end model learning in stochastic optimization." In *Advances in Neural Information Processing Systems*, pp. 5484-5494. 2017.

Clarifying questions: One thing that wasn't entirely clear to me was how the final predicted action/task (i.e. the predicted energy that should produced) accounted for the randomness and variability? More specifically, in one of the figures it showed the energy demand forecast for the next 24 hours with the variance and then the model showed the suggested energy production which was slightly more than the energy demand forecast but within the variance.

Attempted answer: I think the answer for this question is mentioned in the first paragraph of section 4.2. they used a quadratic regularization term in the functions to prevent the coefficients of the generation function from taking too extreme values. Since the weights are slightly less affected by the extreme values in the data the prediction model was able to predict within the variance. They also used 7 years worth of data to train the model and 1.5 years worth of data for testing. Since they are using a very big dataset to train, it is safe to assume that there will be very few outliers and due to their small numbers they will have low affect on the model.

2) Investigate and report on your findings with respect to the "substantive critique" of a different one of the classmates in your group.

Paper: Donti, P., Amos, B., & Kolter, J. Z. (2017). Task-based end-to-end model learning in stochastic optimization. In Advances in Neural Information Processing Systems (pp. 5484-5494).

Substantive critique: The dataset that the researchers used only consisted of 213 images (and 521 faces). While the results obtained from this are a nice proof of concept, in order for this system to be used in any meaningful way I would hope to see a far larger dataset

Report: I agree with the part of the critique that says that 213 images is indeed a small dataset. I have dug around and found a similar research paper where they used images from google's open images. Here is the paper: https://homes.luddy.indiana.edu/kapadia/papers/hasan-oakland-20-preprint.pdf From the database they trained their algorithms on nearly 2,000 images and 5,000 faces, or as they call it in the paper "stimuli". Consequently google itself could train these algorithms on the vast dataset of google photos, which consists of personal photos stored on personal drives. However this raises other potential privacy concerns.