Learning Deep Representation from Deep and Heterogenous Data for Traffic Accident Inference
Goals, Inputs/Outputs

- App tracking incident risk
- Legislation?
- GPS Data
- Incident history with location and severity
- Weather
- Driver state of mind
Mobility (Big Data)

- **Phone GPS tracking**
  - Uploads every five minutes
  - 1.7 million people for 6 months
- **Reduction to amount of movement through a space**
  - Grid based on latitude+longitude (about 440x430 meters)
Denoise Autoencoders (Deep Learning)

- Autoencoder predicts incident risk
- Train many autoencoders, then stack them
- Adjust each layer to get a better final result
- Denoise autoencoder: if we are using training set, add noise to each encoder’s input
Results
Limitations

- Human mobility is not sufficient
  - Weather
  - Driver state of mind
  - Civic traffic control structures (e.g. stop signs, speed limit)

- Kind of obvious results?
Questions

- What is the advantage/disadvantage of using denoise autoencoder, compared with autoencoder?
  - Harder to do backpropagation
  - Noise may cover up features
  - End result is better at dealing with noisy data

- Why do they use the layer-wise algorithm to train the model?
Questions (cont.)

• In this paper, authors are using a version of Auto-encoders for traffic risk prediction but the only reason they provide for this choice is that the task is complicated. Do you think why Auto-encoder is a good option for this task? What's the drawback of other methods like CNN?

• Authors are using denoised data for training and they argue this will improve the performance of their model. However, in the real world, there are lots of noise in the input test data. So do you think it is a wise choice to denoise the data and would it make the model more vulnerable in a real-world setting where there are lots of noise and even adversary inputs to the model?
• It seems to me that the "m" parameter, which defines the width of the human mobility matrix to be used, plays a very important role. However, I don't think the authors discuss it very much or even mention the value that they used for their experiments. Can you think of a good way to select the value of m?

• How are predictions made? It's not very clear to me how the last layer is built. From Figure 4 it looks like they predict several values at the same time, are they predicting the risk value for several regions at the same time? Or is it a softmax over all the possible risk values?
Figure 4
In the paper the authors remark that "Human mobility predetermines traffic accident, for crashes occur between moving vehicles and moving people". Does this mean they are only considering accidents between vehicles and humans and not between two vehicles?

It says the human mobility matrix for a particular region should include the information from adjacent regions. Why do they think this is necessary? Is it done simply to reduce noise?