Question 1

a.

Develop a simplistic diagnostic system

- The problem seems to be a binary classification, whether or not, the damper has flaws. The data are images and a power spectrum of their audio responses. There are a few ways to develop this system.

- The first one would be to solely rely on the images and develop a CNN model as the data is labelled this would be trained with a backdrop and the loss function would be cross-entropy.The second way would also be a CNN model however the last convolutional + pooling layer you concatenate it with the vector and put it through fully connected layers to do the classification The third one would be to use a sparse autoencoder or Deep Belief Network to learn representations and then feed it into the CNN after the last Conv + pool layer to do classification.

- How deep this model would need to be is up for model selection. The modelection is done with Area under curve score. To do model selection, divide the model into train, validation and test datasets.

They wonder if one network could be used to produce a reliable .. for all cars:

- For this, train one model for all the cars then compare it with models trained on specific cars. Test these models against each other to see how well the first model generalizes. This should be done on many sites then their results should be aggregated.

If it suffices to rely only on the frequency responses:

- An MLP could be used for this and compare it with the multimodal model. The MLP model would take the raw data or the representations learned from DBN networks.

b.

To check if the “site matters” for the car. You take the same amount of tests of all the sites and aggregate them to compare them to each other.

c.

Network: Seems to be a clustering problem, SOM networks are good for this with output grid at 3D. Training done with competative learning, preferably leaky learning. The neighbourhood size depends on the amount of distinct groups however you could also start with a high neighbourhood size and adjust it as it. To compare similarity, use cosine, manhattan distance in 3D or euclidean distance. Compare all the groupings and check wich grouping is closest to the good damper cluster. Train a SOM on all car sites and one individually for all car sites and compare how well the first SOM does against the other

d. Extend the output. Previously it was a binary output, now it can be extended into good, errors and not production ready. Makes it easier to distinguish between good dampers and flawed damper categories.

Question 2

a)

p * w = [3,1,1, -1, -1, 1]

[1,1,1,-1,-1,1]

second iteration:
[1,1,1,-1,-1,1]

Seems to be stable. The only difference is that the first bit is flipped. Thereby P is a noisy version

b) The nodes in RBM are stochastic. Explain how they are stochastic and what role the stochasticity plays in the function of the network

Question 3

A) Neural network: MLP with feed-forward architecture. Seems to be a regression/time series forecasting problem. Input is the data and the output is forecasted water speed, number of waves per unit area, size of waves etcetera depending on the input. The MLP is trained with back-prop and Mean squared loss for the regression tasks. The hidden layers and the nodes inside the hidden layers are to be selected in the model selection process

B) Neural network: MLP With feed-forward architecture. Seems to be a classification task whether or not with the classes either normal or abnormal states. If failure is low then it can be seen as anomaly detection. Trained with back-prop to recognize input correlations correlated with system failure. Input layer would correspond to the number of sensors and output would depend on how many output classes. Sigmoidar activation function equal to classes. Build models trained on a single sensor versus a model trained on all sensors and check which one performs better on a test sate and through cross-validation. If multivariate model performs significantly better then the hypothesis hold weight else it doesn't. The aspect of data which could class problems & challenges are overrepresentation of a single class. Also risk of overfitting due to curse of dimensionality.To chek the other hypothesis, train a MLP on time-delay reåresentations of sensor readout and one with current input states. Compare them with cross-validation and conclude.

C) .
D) .

Question 4

Question 5 - 0.5 points

Problem: Seems to be a variant of the Travelling Salesman problem
Network: Hopfield or SOM however Hopfield is preferred due to it’s ability to solve opåtmization problems
Topology: The Hopfield dimension would be n^2 where is the amount of paths. In this case the number of paths are the number of hotels. Let 1 indicate a hotel with guests and -1 hotel with no guests. No training conducted however initalized to minimize energy. Set penalty for increased total distance

Question 6 → Difficult question thoguh

- Network: This is a classification problem with non-linear labelled data. RBF networks. A shallow MLP is preferred to minize overfitting. Learning algorithm for MLP is backprop with regularisation to prevent overfitting, competative learning for RBF network with least squares at the end model. Model selection through cross validation. To minize overfitting, regularisation techniques should be used, this is a concern as there is a lot of dimensionality in the data.

- To interpret an output, have 5 classes in the output where each corresponds to a triage category. The highest priority with a the highest value and the lowest with the lowest value. Measure performance with true positives on the priority patients. The challenge the curse of dimensionality as we have a lot of dimensions (56) however roughly 7-9 questions have been answered by 36 patients where as 12% of the correspond to the highest priority targets. See if possible to reduce dimensionality to risk overfitting and care to the ones with highest priority.

- Encoding questions 1-7 binary or compress them hard. Split into 300 training, 100 validation and 100 testing.

- Extended questionaries adds more dimensionality however given sparse representations in the data and missing values, this makes the model more worser to generalize and prone to overfitting. The most important patient category is the patients with the priority.