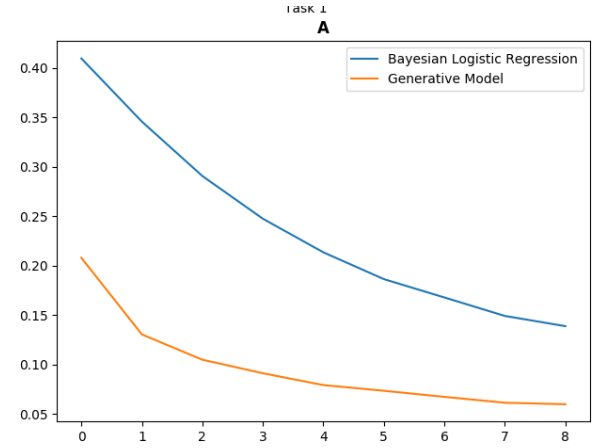
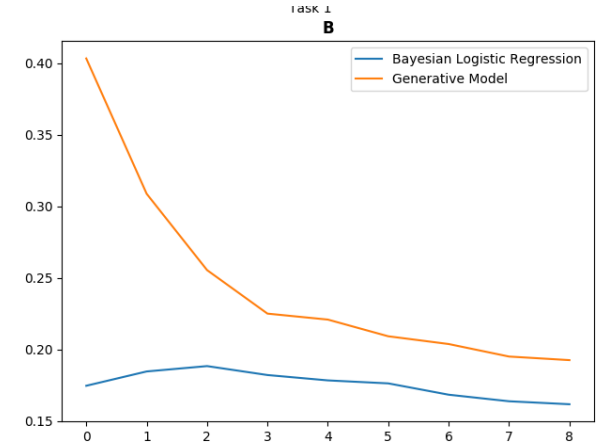
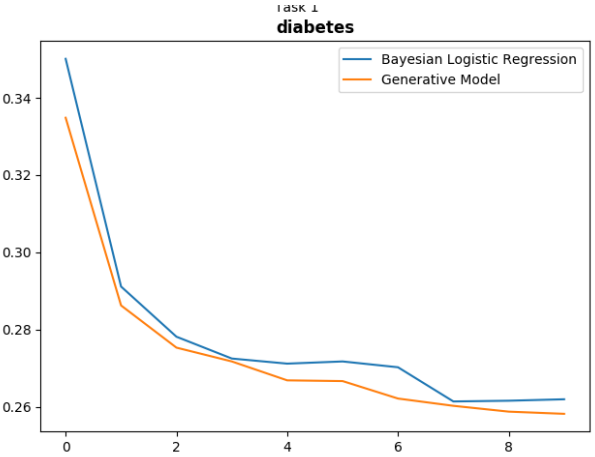
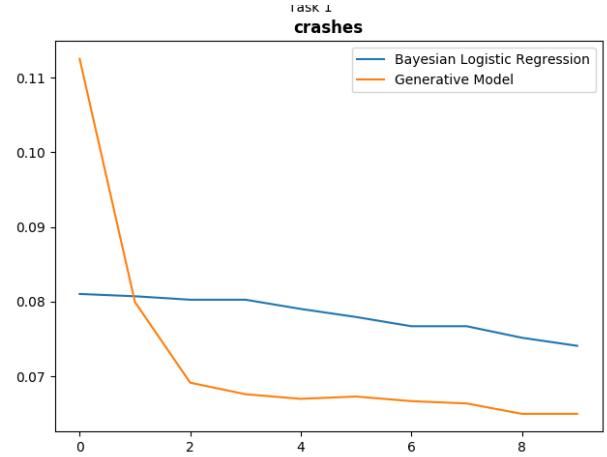
Alexa Bosworth
Programming Project 3
COMP 136

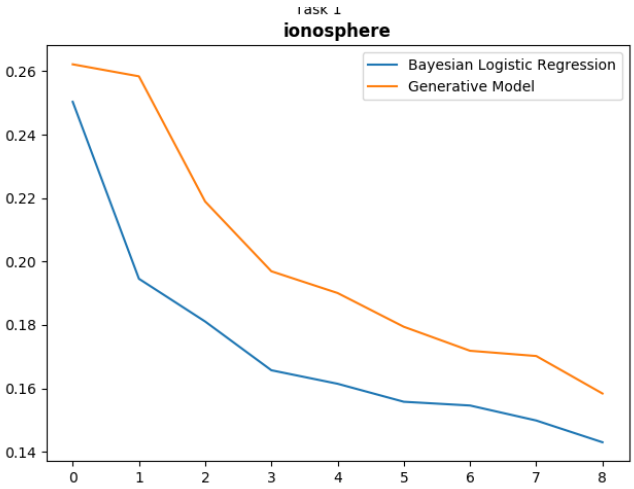
Task 1





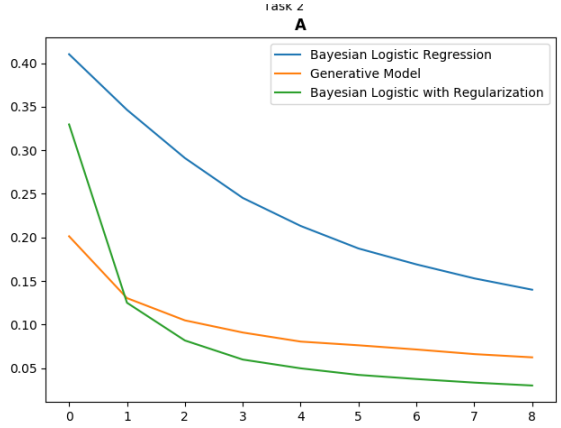


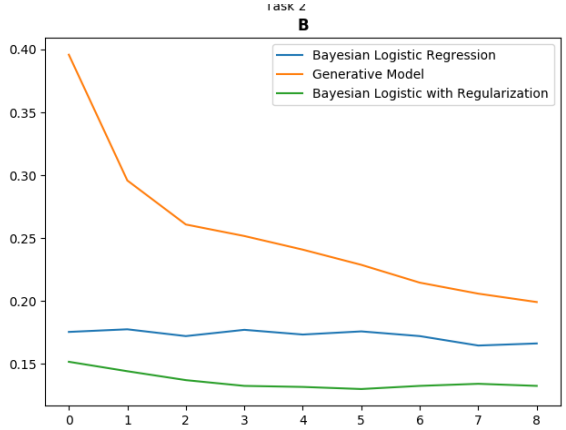


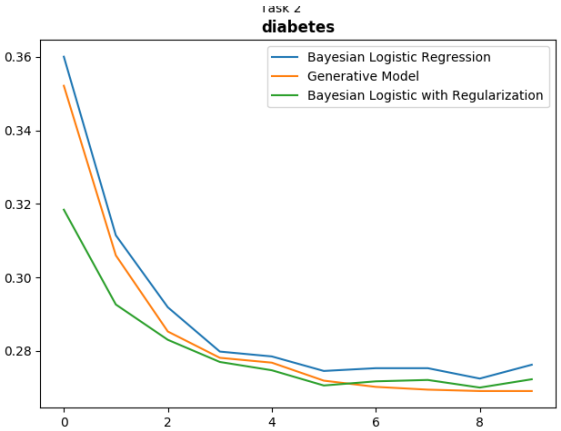


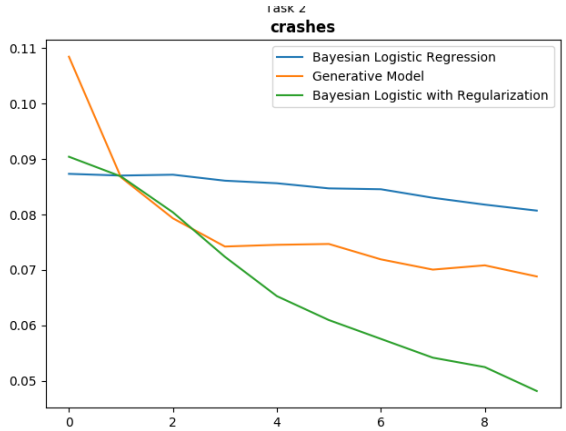
From what we know about the A dataset, it is understandable that the logistic regression does not learn the model as well as the generative model, as the data is uniformly distributed but the logistic regression model utilizes a normal distribution prior and likelihood.
Dataset B is composed of multiple Gaussians, thus the logistic regression model is better suited model than dataset A.
The diabetes dataset show that both models perform creatively the same.
The crashes dataset achieves high accuracy from both models. Logistic Regression appears to be able to learn the model quickly but accuracies converge almost instantly.
The ionosphere dataset shows that the logistic regression model outperforms the generative model, but if more training was allowed, it is possible that both model accuracies could converge.

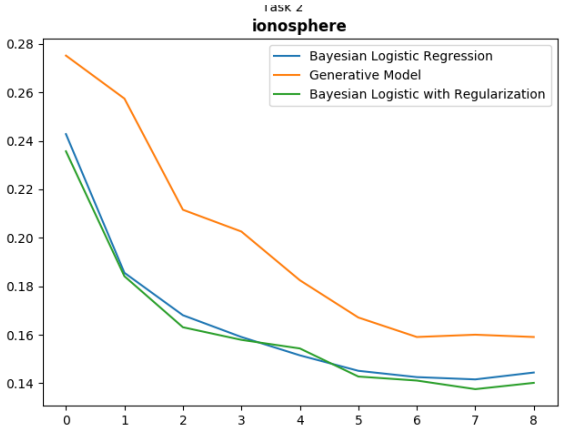
Task 2





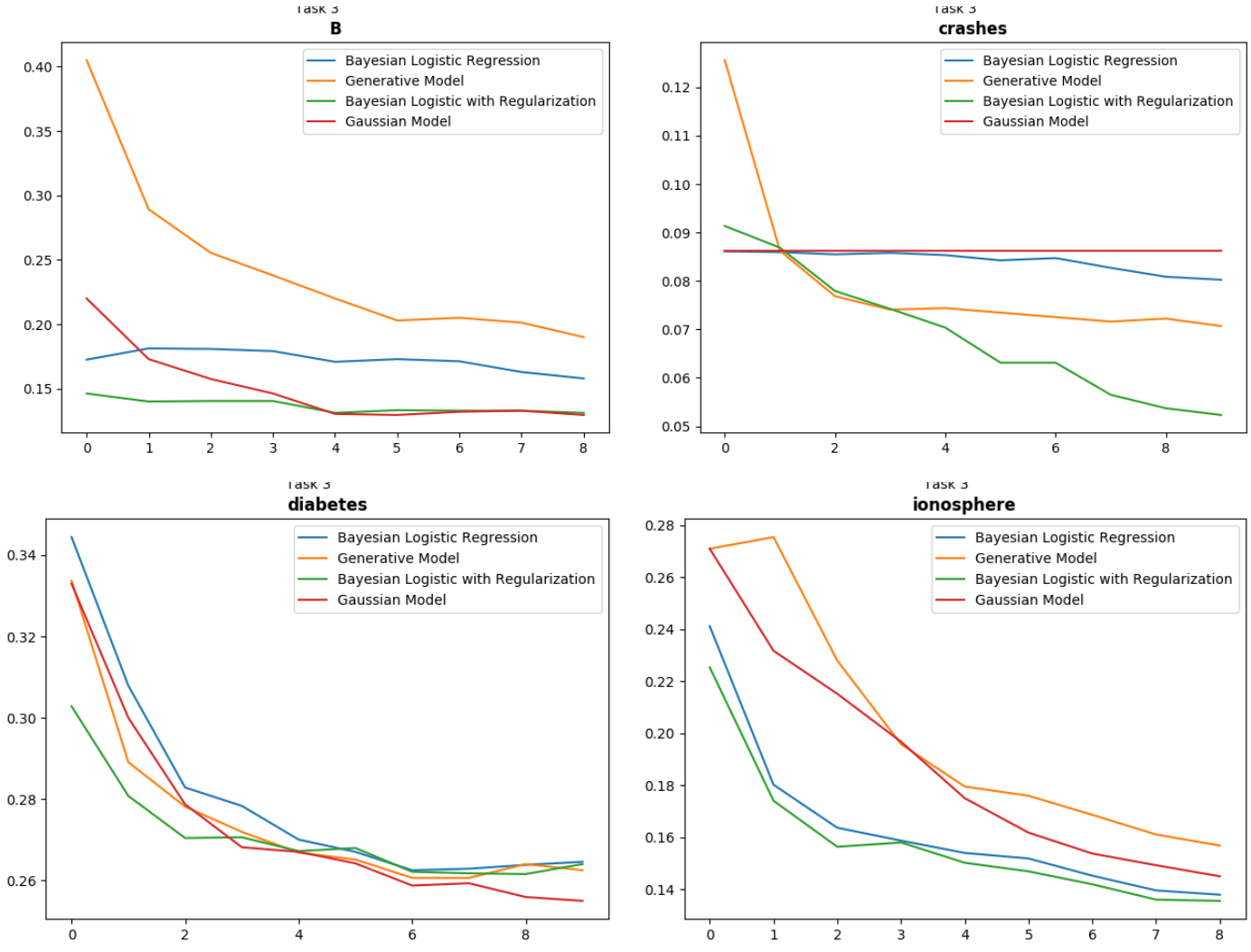






By optimizing for hyperparameter a, we see that bayesian logistic regression with regularization outperforms in every dataset, to varying degrees. In fact, it outperforms the generative model as well. Therefore, on average, the regularized bayesian logistic regression converges faster and allows the model to generalize better over the held-out set. The function of regularization is to help a model generalize better over unseen data by preventing the model from overfitting. Therefore, it only follows that a model with an optimum* regularization term value will outperform a model without regularization which tends to overfit.

Task 3



It is worth noting that Gaussian Processes are computationally expensive and given the results in comparison to the models we have developed, the GP model does not significantly outperform the other models (except of course in the diabetes dataset, but not significantly). Given this, I would not chose GP for any of the datasets (except maybe possibly diabetes). I find the GP curve (or lack thereof) for crashes to be puzzling. My guess is that GP cannot improve on its initial model and so it didn't optimize further with more training data.