ADNI Progress report

Devendra Goyal
Uniqname: devendra

November 7, 2014

Part I

Using HMMs to predict disease progression

1 Statistics about the data

There are several choices here for the dataset to be used:

- One of MRI/PET/CSF only
- Some combination of 2/3 available modalities
- Use all three modalities
- Add demographics data to any of the above options

For now, I have decided to use only the FDG-PET data for the HMM problem. Because,

- The PET data is the cleanest and most easy dataset to process.
- The feature space is relatively low dimensional. It is the

  $$mean, median, mode, min, max, stdev$$

  for 5 expert defined regions in the brain - making it a total of 30 features for each patient.

Part II

Establishing the baseline

2 Statistics about the Data

The results shown in this document are based on the ADNI-1 cohort. There are several pros and cons of this decision:

2.1 Pros

- All the pre-processed features uploaded on the ADNI website operate on the entire cohort, thus there is homogeneity in terms of the features available for each patient across each modailty. MRI images for ADNI-GO/2 have been processed using different versions of the same software, and are also collected using a higher resolution MRI machine. While this is not a dealbreaker per se, it will require significantly more work to process all the images using the same software and generate similar features.
- Most of the recent literature (5 years) relies on data from ADNI1 only to report results. This will give us a chance to compare our results directly with some of these reported results.
- ADNI-1 is the best dataset to track patients longitudinally, as it was started in 2004 and we have about 8 years' worth of data for all MCI and AD patients that the protocol chose to follow (more on this later).

2.2 Cons

- ADNI1 has CSF data available for only about 20% of the patients. This number will become even smaller when we look at the number of patients that have data available for all 3 modalities.

Below is a chart summarizing the study design:

The table below summarizes the number for the **ADNI1** cohort only, for the MRI and PET modalities.

	Normal	EMCI	MCI	LMCI	AD	MRI	fMRI	DTI	FDG	AV45	PIB	Biosamples
ADNI 1	200	–	400	–	200	✓			✓		✓	✓
ADNI GO	↓	200	↓	–	–	✓	✓	✓	✓	✓		✓
ADNI 2	150	150	↓	150	200	✓	✓	✓	✓	✓		✓

Figure 1: ADNI study design

	NL	MCI			AD
		MCI-c	MCI-nc	MCI-rev	
FDG-PET	102	94	96	13	97
MRI(clean)	180	133	132	10	123
MRI(complete)	229	196	183	18	192
FDG+MRI(clean)	79	66	68	9	62
FDG+MRI(complete)	102	94	96	13	97

3 SVMs on classification task

3.1 NL vs AD

3.1.a MRI

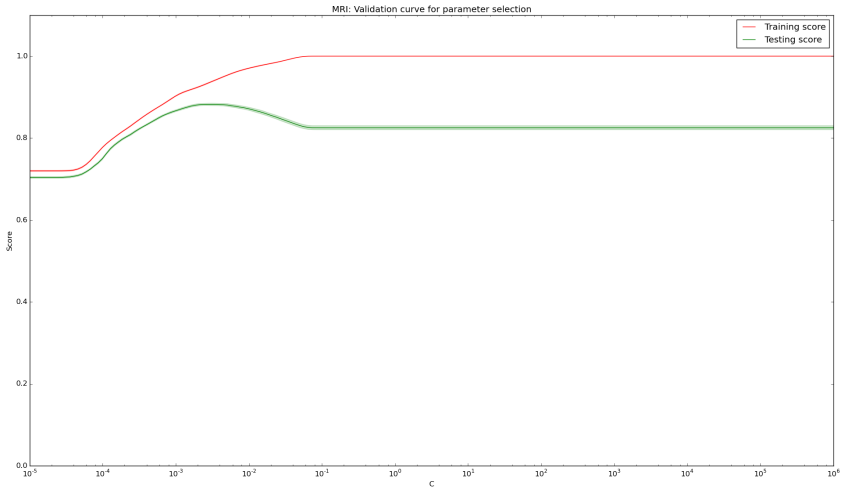

Figure 2: Hyperparameter search for MRI images optimizing for classification accuracy

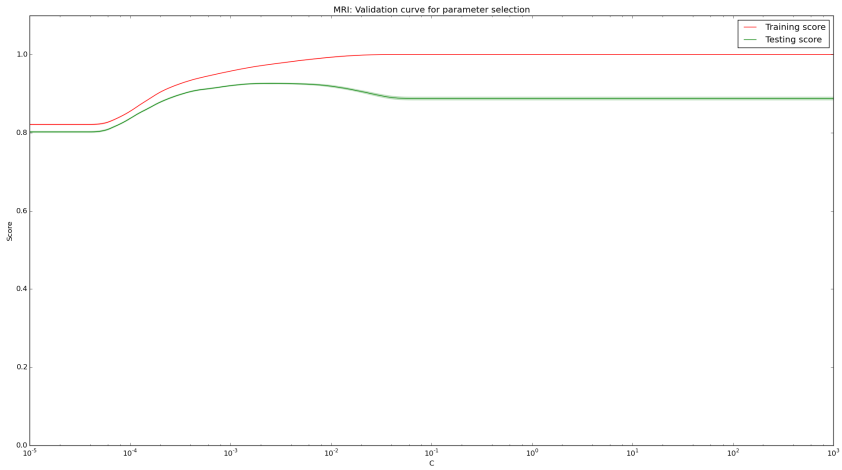

Figure 3: Hyperparameter search for MRI images optimizing for AUROC

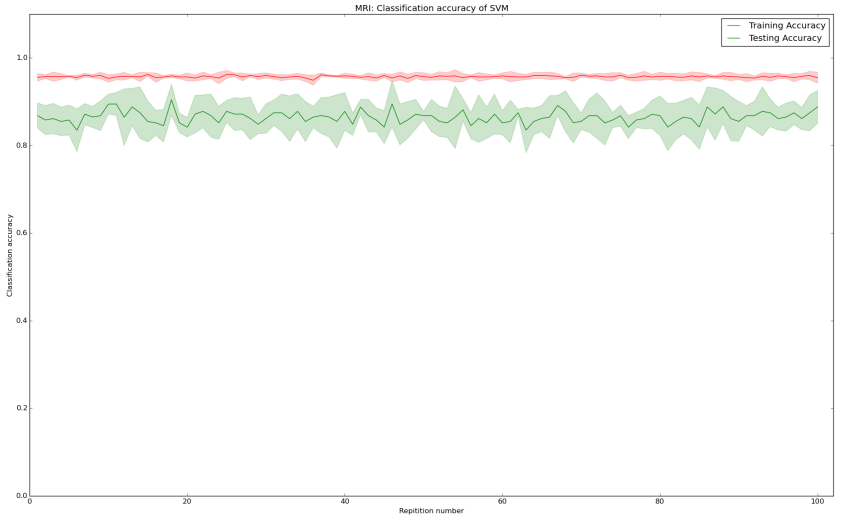

Figure 4: Classification accuracy (test set) for MRI images

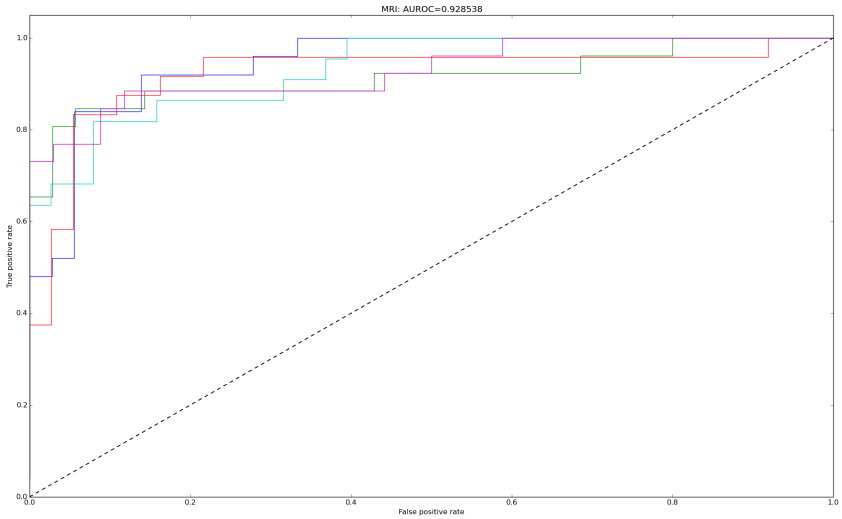

Figure 5: AUROC for MRI images (test set)

3.1.b PET

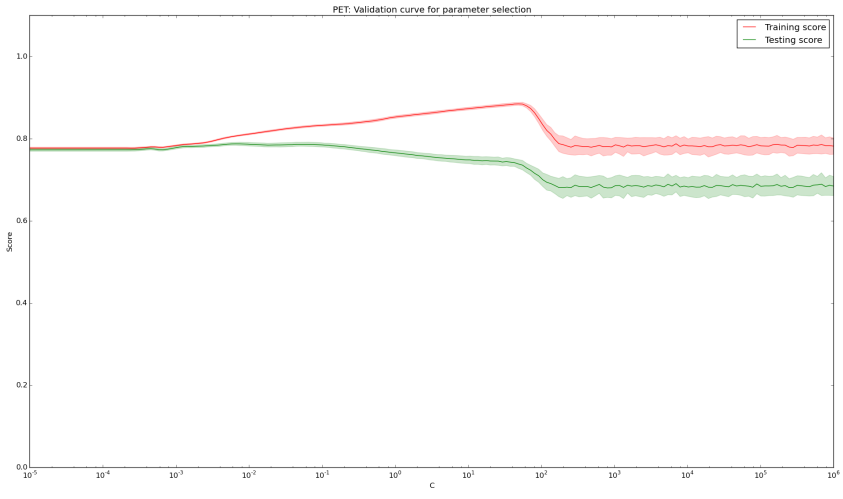

Figure 6: Hyperparameter search for PET images optimizing for classification accuracy

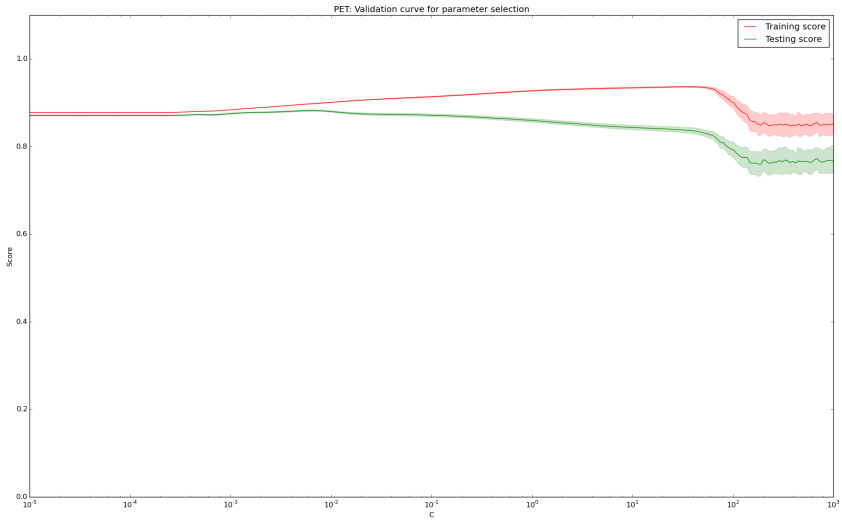

Figure 7: Hyperparameter search for PET images optimizing for AUROC

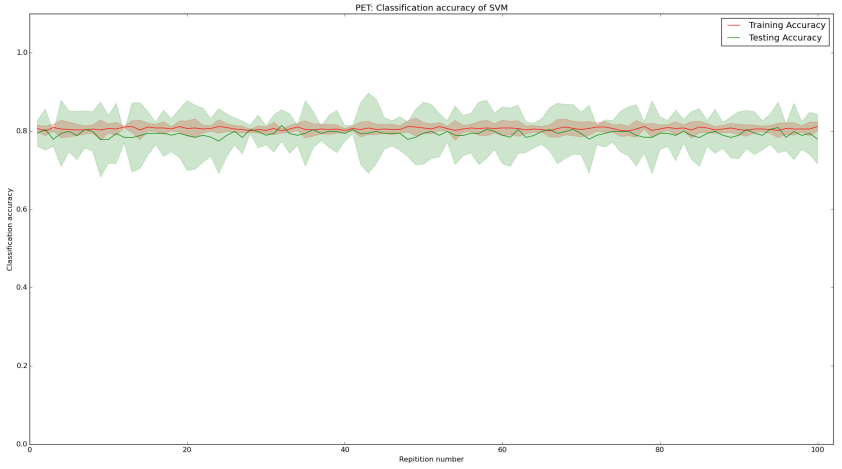

Figure 8: Classification accuracy (test set) for PET images

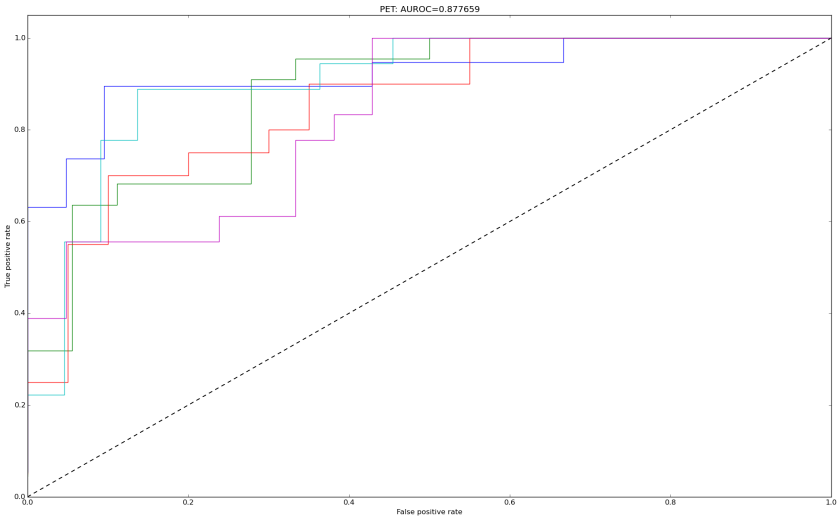

Figure 9: AUROC for PET images (test set)

3.1.c CONCAT

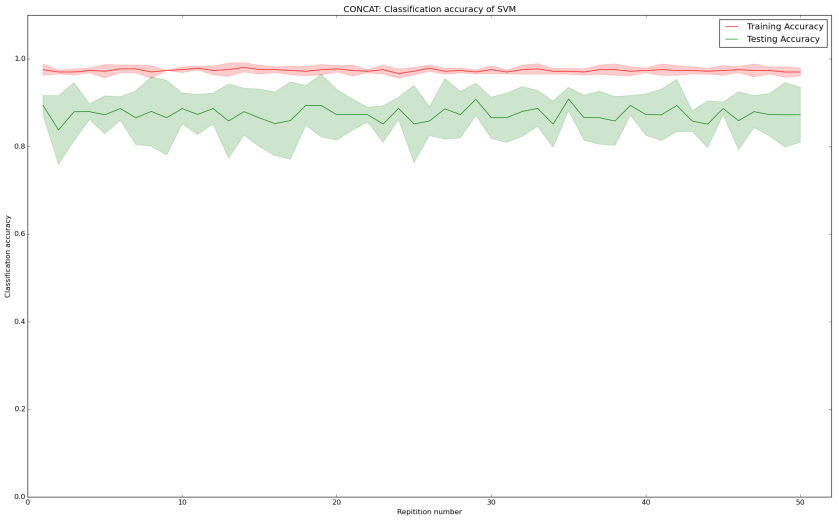

Figure 10: Classification accuracy (test set) for PET images

3.2 NL vs MCI

3.2.a MRI

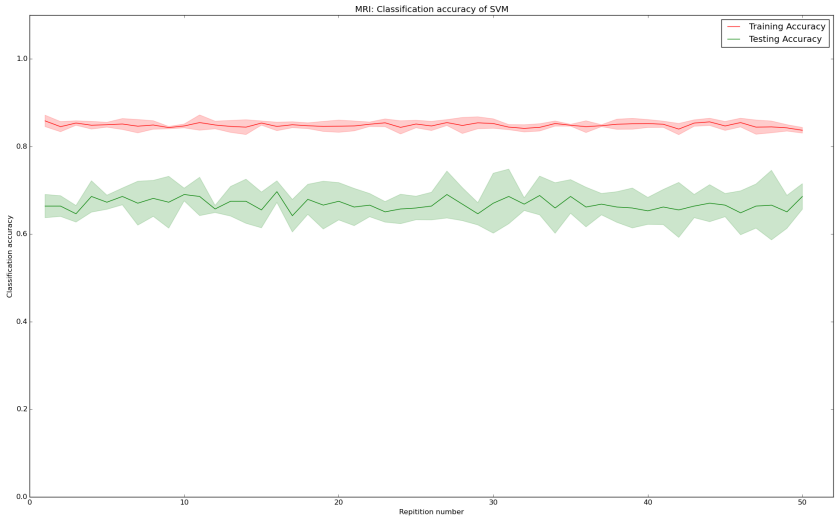

Figure 11: Classification accuracy (test set) for MRI images

3.2.b PET

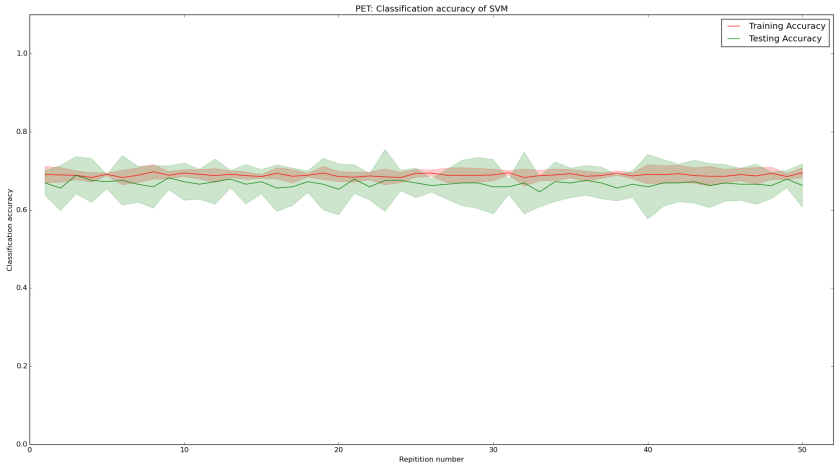

Figure 12: Classification accuracy (test set) for MRI images

3.2.c CONCAT

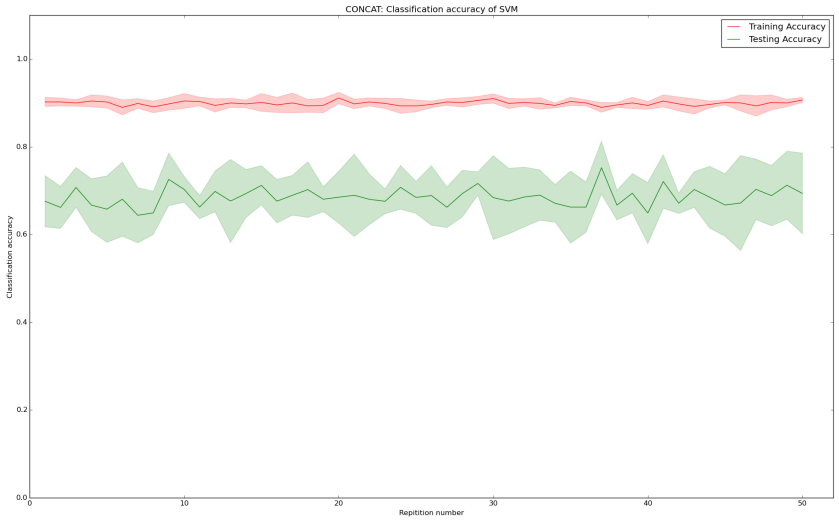

Figure 13: Classification accuracy (test set) for MRI images