

All code to reproduce the submitted results is available in a git repository `git@bitbucket.org:alex-hh/cagimpra.git` which can be made available on request.

Features

We tested different types of feature to assess their predictive power.

Conservation

We used *phyloP*, *phastCons*, *GerpN* and *GerpRS* base-pair resolved conservation scores.

DNase hypersensitivity

For each regulatory element we identified the closest matching ENCODE cell type for which there is a DNase-hypersensitivity track. We downloaded the tracks and created 41-dimensional features consisting of the signal at each base-pair flanked by the signal at the bases 20bp on either side.

DeepSea neural networks

We trained several different neural network architectures on the genomic prediction benchmark detailed in the original DeepSea paper. We evaluated these networks on a region surrounding each variant twice, once with the reference allele and once with the alternate allele. Features were generated as the difference in activations between the two evaluations of internal and output layers of the networks.

Region identity

We one-hot encoded the identity of region to use as a feature.

Substitution

We one-hot encoded the reference to alternate allele substitution as a feature.

Inference

We used three different gradient boosting algorithms: *XGBoost*, *CatBoost* and *LightGBM*. We used the gradient boosting algorithms to regress the **Value** and **Confidence**, we also used them as classifiers to predict the **Direction**. We tested different subsets of Features in a 5-fold cross-validation set up to identify the best features. Models were assessed by the cross-validated one against many area-under-precision-recall-curve (AUPRC). We stacked models in that we used

the prediction of cross-validated regression models of the **Value** and **Confidence** as features in the classification model for the **Direction**.

To estimate the standard error of the **Confidence** we trained an ensemble of regression models and used the standard deviation of the predictions to estimate the standard error.