Neural Net

The neural net classifier was constructed with four layers of nodes, the input layer, two filter layers, and the output layer. The input and filter layers contained 1100 nodes and the output layer contained 1106 nodes for each of the go terms. These parameters were found to achieve the best results on the testing data through hyper parameter tuning. The average AUROC was 0.585 and there were no significant changes in AUROC with increase in support. The weighted F1 score was chosen to reflect the severe class imbalance in the positive go annotations. The weighted F1 scores indicate that the neural net contained moderate predictive power given the small set of positive instances. The binary F1 scores also indicate a trend of increasing F1 score with increasing number of positive go annotations.

The go terms with a binary F1 score greater than 0.1 with definitions on what the go terms are associated with from chat gpt

0.10030090270812438 GO:0006397 - mRNA processing: This term refers to the various processes involved in the conversion of precursor messenger RNA (pre-mRNA) into mature mRNA, including capping, splicing, and polyadenylation.

0.10122699386503069 GO:0045786 - Negative regulation of cell cycle: This term refers to the processes that prevent or slow down cell cycle progression, including the regulation of cyclins, CDKs, and checkpoint controls.

0.11904761904761903 GO:0038061 - NAD(P)+ binding: This term refers to the binding of nicotinamide adenine dinucleotide (NAD+) or nicotinamide adenine dinucleotide phosphate (NADP+) to a protein or enzyme.

0.1241565452091768 GO:0006281 - DNA repair: This term refers to the various mechanisms by which cells repair damaged DNA, including nucleotide excision repair, base excision repair, and double-strand break repair.

0.13197969543147206 GO:0070125 - Mitotic spindle checkpoint: This term refers to the process that ensures the proper alignment of chromosomes during mitosis, and the arrest of the cell cycle until all chromosomes are correctly aligned.

0.15204678362573099 GO:0070126 - Mitotic spindle organization: This term refers to the processes involved in the assembly and organization of the mitotic spindle, including the formation of microtubules and the positioning of centrosomes.

0.15384615384615385 GO:1900026 - Positive regulation of rRNA binding: This term refers to the processes that increase the binding of ribosomal RNA (rRNA) to ribosomes or other RNA-binding proteins.

0.16 GO:0050918 - Positive chemotaxis: This term refers to the movement of cells or organisms toward a chemical stimulus, such as a nutrient or signaling molecule.

0.25 GO:0042594 - Response to starvation: This term refers to the physiological and biochemical changes that occur in response to a lack of nutrients or energy, including changes in gene expression, metabolism, and behavior.

0.3333333333333333 GO:0007029 - Endoplasmic reticulum organization: This term refers to the processes involved in the formation, maintenance, and organization of the endoplasmic reticulum, including membrane biogenesis, protein targeting, and quality control.

0.7999999999999999 GO:0006120 - Mitochondrial electron transport, NADH to ubiquinone: This term refers to the transfer of electrons from NADH to ubiquinone (coenzyme Q) during oxidative phosphorylation in mitochondria.

0.9411764705882353 GO:0032981 - Mitochondrial respiratory chain complex I assembly: This term refers to the processes involved in the assembly and biogenesis of complex I (NADH-ubiquinone oxidoreductase) in the mitochondrial electron transport chain.

