Deep Learning

Next Generation Gene Analysis

The Goal

is to identify exitron splicing events and patterns which disrupt pre-mRNA relevant to cancer. Those event and patterns will lead scientist to identify novel biomarkers for cancer diagnosis.

The Problem

In eukaryotic cells, protein coding information intrinsic to exons is interrupted by introns. Introns have to be removed and exons spliced together to provide a template for protein synthesis, mature RNA (mRNA). Exons and introns or their parts can be differentially included in mRNA by so called alternative splicing (AS), thus resulting in multiple protein variants with different fates and functions from a single gene.

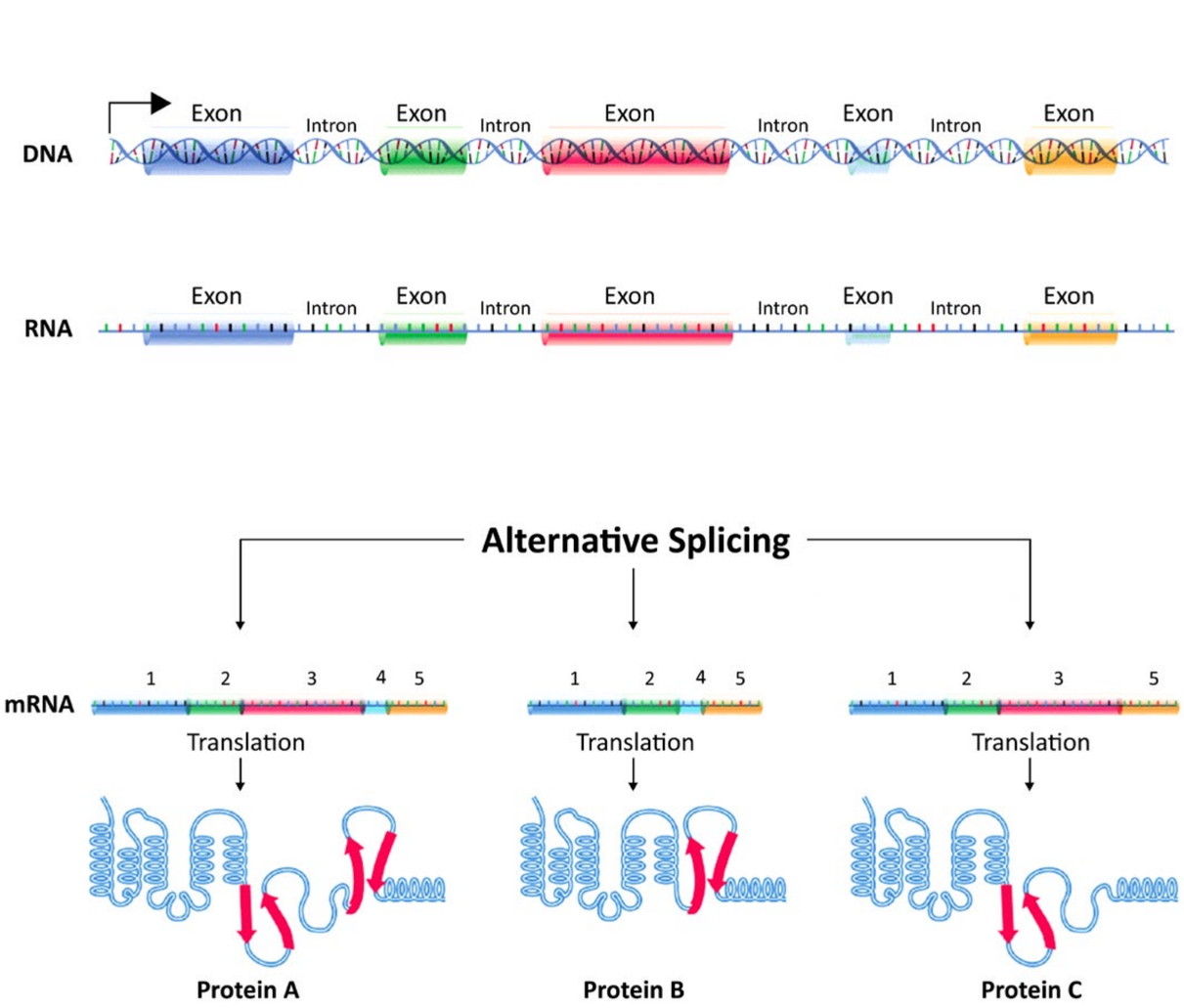
AS is a mechanism to increase the density of information encoded by eukaryotic genomes as about 95% of humans are alternatively spliced (Chamala 2015). The genome has 33.000 genes, but through alternative splicing many hundred thousand proteins can be generated – so AS is the key to understand the genome.

The Way

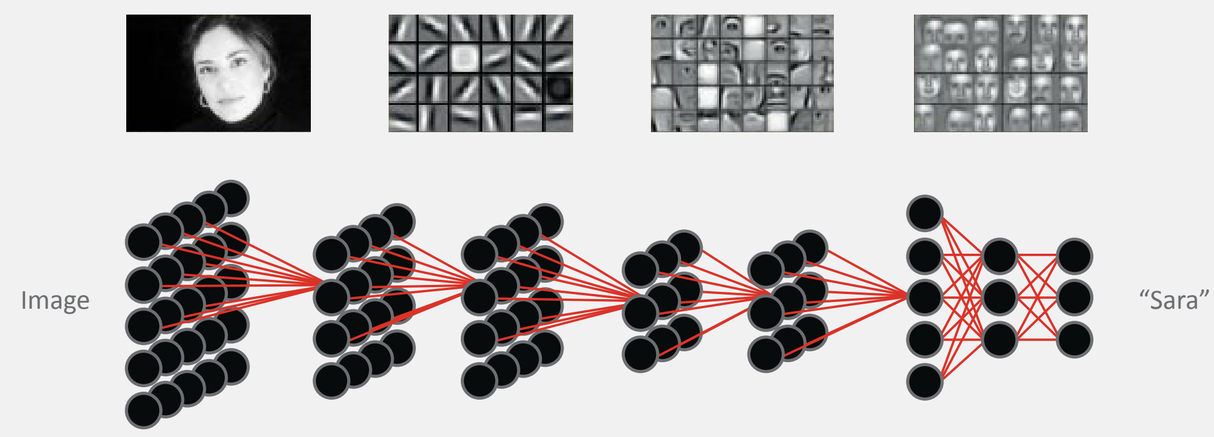
is to use machine learning and deep belief nets to identify disrupted pre-mRNA caused by exitron splicing events. A critical task is to prepare the training and evaluation data sets.

Fascinating about deep belief nets is their remarkable ability to generalize beyond the training examples. The hidden layers really specialize on the detection of certain patterns.

Splicing



Deep Learning



Medizin- und Bioinformatik

we are Catalysts
software is our passion