CCDS Radiology Analytics
Facilitating Research and Diagnosis

Problem Statement: The Center for Clinical Data Science is a new center in Boston created to promote, develop and commercialize artificial intelligence for healthcare. They are a growing team consisting of researchers, healthcare professionals, and entrepreneurs.

CCDS is working with radiologists to develop a tool that enables the search and extraction of information from radiological medical reports. When radiologists read an imaging study, they report their findings in plain text. The unstructured information contained in large collections of reports contains the synthesis of the clinical knowledge of the radiologists and of their expertise; CCDS have created a search system based on Regex and on the SpaCy NLP library that enables finding reports in a database that mention a certain medical condition. Once a report is found in this way, it is manually annotated by a domain expect to indicate in a structured form the properties of the medical condition described in the report. Once a large enough number of reports has been annotated, the annotations can be used to create sets of images and train visual neural nets that learn to perform image based diagnosis. Below is an example of the output of a visual neural network that has learned to detect hemorrhage.

Project Goal: The goal of this project is to enable more efficient search, extraction of medical information, and annotation of these medical reports. To do this, CCDS envisions two tools.

The vision for the first tool is to take a new report (that has still not been annotated) and to probabilistically classify it along several predefined dimensions. This is shown in the example below, where a medical report is classified based on three categories, presence, location and type of hemorrhage. CCDS envisions training a neural net to achieve this. Such a tool would facilitate faster diagnosis and also make the process of subsequent report annotation faster.
The vision for the second tool is to take a search term and rank all existing reports (both annotated and unannotated) based on relevance. As shown in the example below, the tool takes search string consisting of a particular hemorrhage condition and returns ranked reports. CCDS envisions training a neural network to achieve this.