# Medical Visual Question Answering

*Weekly Report 1*

Dataset : VQA-Med @ ImageClef 2019

Visit this link for an interactive chart showing the different categories in the dataset along with their frequent answers.
Here is the jupyter notebook with all the code and instructions to run.

## Highlights:

- Exploration of Question and Answer part of the dataset
- **Clinical word embeddings (pre-trained on PubMed and PMC articles)**
- t-SNE and Cosine-Similarity Plots
- A baseline question classifier using highest cosine-similarity. **Accuracy : ~80%**

## The Task

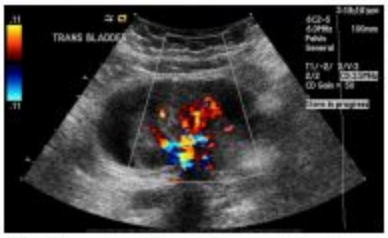
Since the release of the first VQA dataset in 2014, additional datasets have been released and many algorithms have been proposed. The dataset examined here is VQA-Med, based on images from the MedPix database with only the cases where the diagnosis was made based on the image.

Although it has been made sure that the current task would not have answers requiring external domain knowledge, such a dataset is proposed to be introduced in future ImageClef VQA tasks. It can prove to be useful in practical applications with open-ended questions to first learn/use pre-trained embeddings for the medical entities present in the textual data.

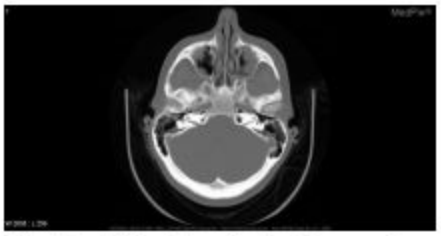
In this weekly report, a basic exploration of the question answer data has been performed. Clinical embeddings pre-trained on PubMed & PMC full texts were analysed. A basic classifier which calculates cosine-similarity of embeddings to predict the question category showed an accuracy of ~80% on the training dataset's unique questions.

# Dataset

The 4 categories the dataset has questions on are : Modality, Plane, Organ System, Abnormalities. A few examples from test set are shown below:-



(a) **Q**: what imaging method was used? **A**: us-d - doppler ultrasound



(b) **Q**: which plane is the image shown in? **A**: axial

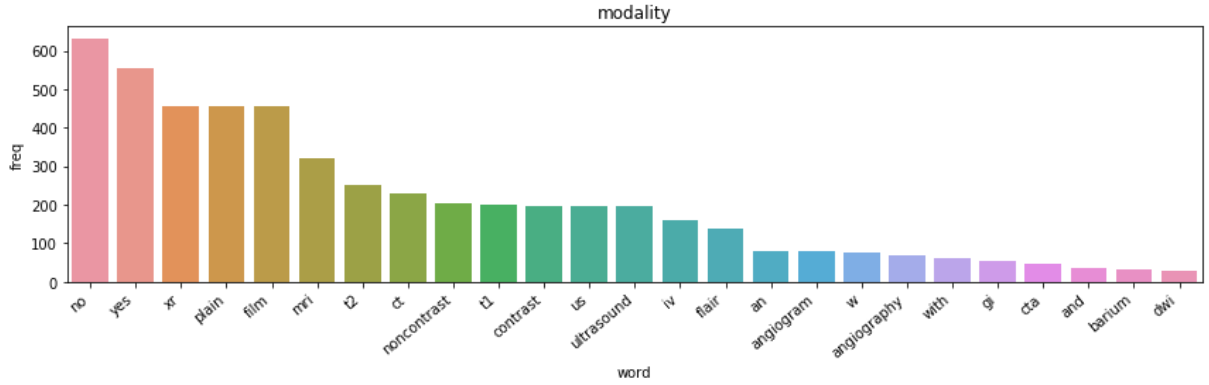
(c) **Q**: is this a contrast or non-contrast ct? **A**: contrast

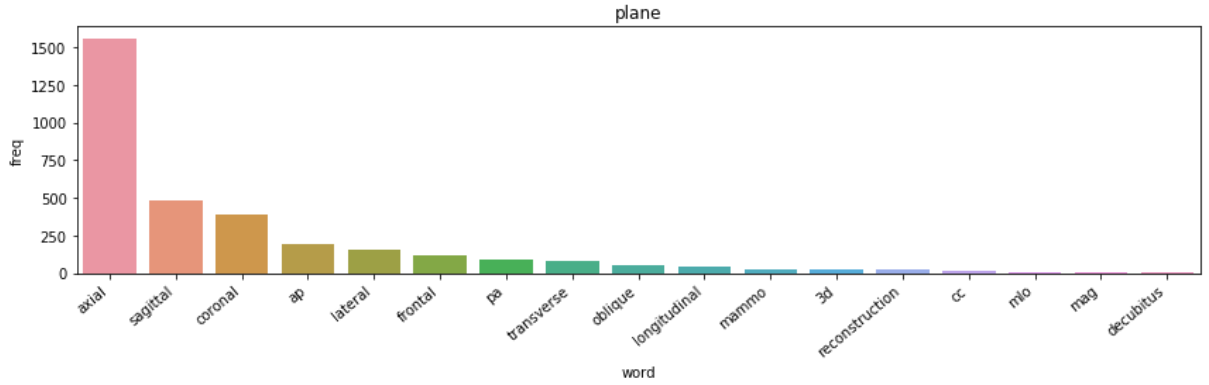
(d) **Q**: what plane is this? **A**: lateral

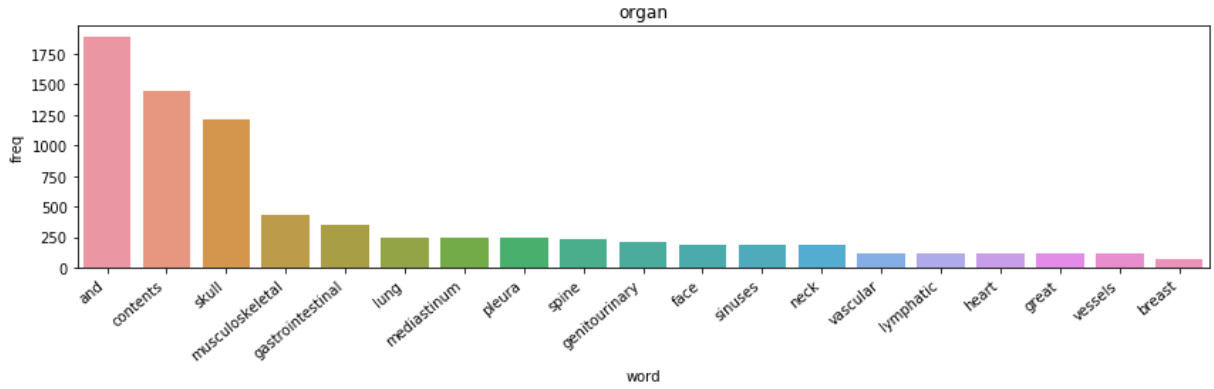
## Length of Q/A

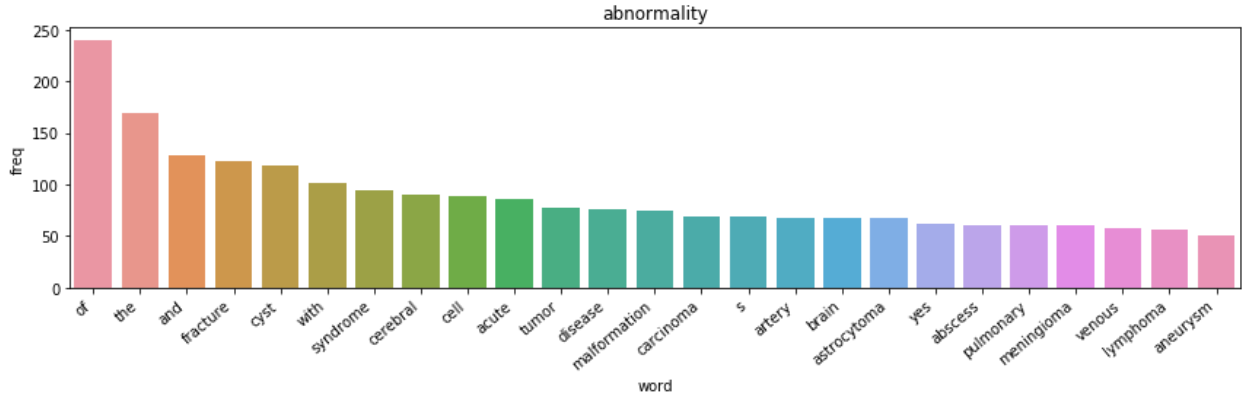
| Category | Av. Words/Question | Av. Words/Answer |
|---|---|---|
| All | 7.9 | 2.1 |
| Modality | 7.7 | 1.8 |
| Plane | 7.2 | 1.0 |
| Organ | 8.4 | 2.4 |
| Abnormality | 8.1 | 3.4 |

# Frequent Words in Answers



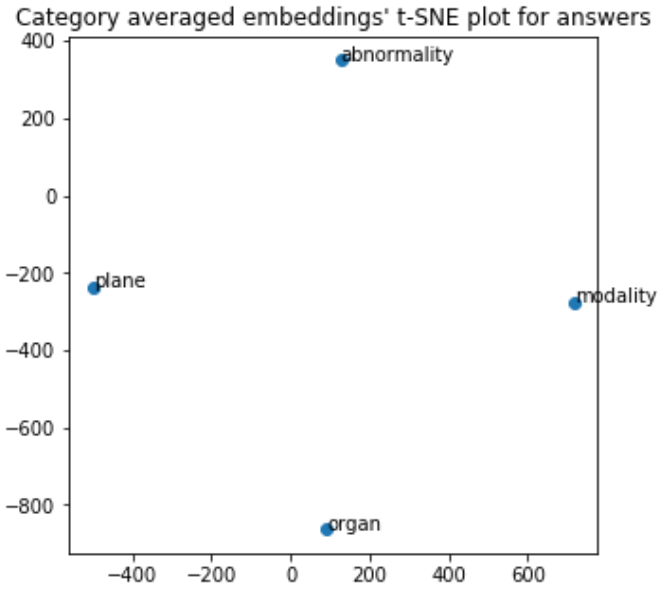



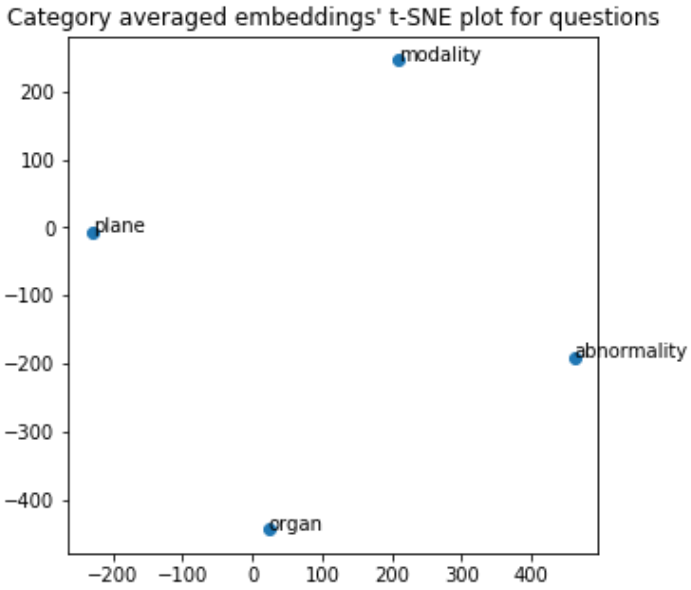



# Clinical Word Embeddings

Using 200 Dimensional Word2Vec style Embeddings pre-trained on **PubMed and PMC Full Texts**. Released by Pyysalo, Sampo et al. “Distributional Semantics Resources for Biomedical Text Processing.” (2013) with a total of 5B+ tokens. (Link)


Plot shows **consistency with the relatedness of concepts** such as:- yes and no lying opposite to each other, planes confined to a small area, t1 and t2 near each other etc.



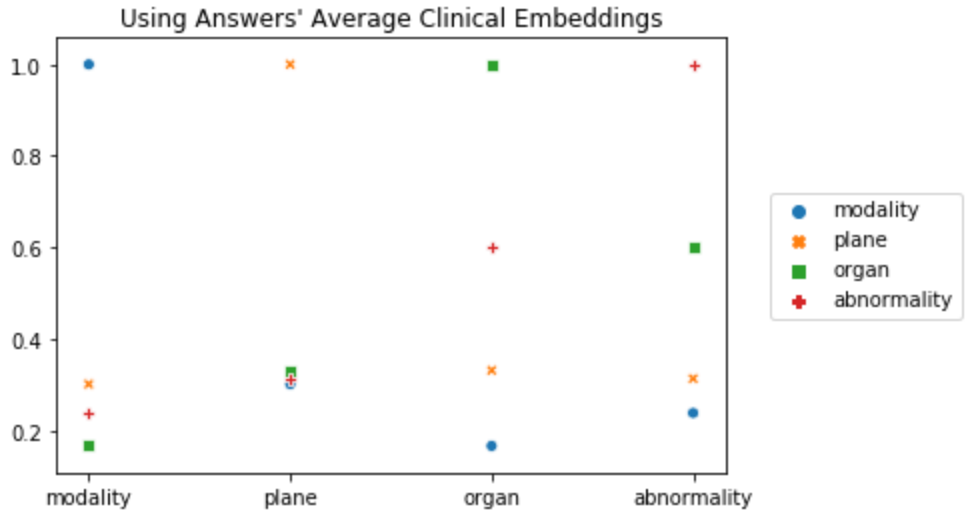

## Similarity scores of clinical embeddings

| *Word 1* | *Word 2* | *Similarity Score* |
|---|---|---|
| medicine | pediatrics | 0.77 |
| infarction | myocardial | 0.73 |
| cancer | carcinoma | 0.67 |
| fracture | osteoporosis | 0.57 |
| longitudinal | plane | 0.4 |
| modality | mri | 0.37 |
| cannula | insert | 0.3 |
| crohn | ileum | 0.25 |
| angina | osteoporosis | 0.2 |
| contrast | noncontrast | 0.12 |
| artery | mri | 0.09 |
| skull | sunlight | 0.07 |
| cannula | show | -0.08 |

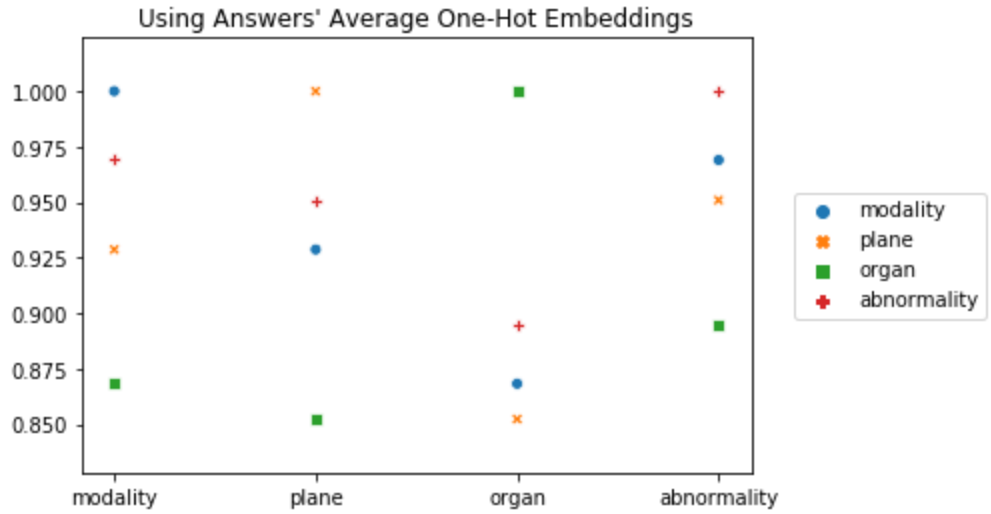
## Top 10 Similar Words (and similarity scores)

| failure | plane | imaging | musculoskeletal |
|---|---|---|---|
| ('congestive', 0.7064043283462524) | ('planes', 0.8751378059387207) | ('MRI', 0.8483814001083374) | ('work-related', 0.6567299962043762) |
| ('decompensated', 0.6755109429359436) | ('perpendicular', 0.8055003881454468) | ('MR', 0.8094216585159302) | ('rheumatologic', 0.6503987312316895) |
| ('CHF', 0.661030650138855) | ('z-axis', 0.7746583223342896) | ('contrast-enhanced', 0.7706788182258606) | ('low-back', 0.6341665387153625) |
| ('end-stage', 0.6437994837760925) | ('tilted', 0.7433577179908752) | ('gadolinium-enhanced', 0.753862738609314) | ('genitourinary', 0.6101983189582825) |
| ('decompensation', 0.6201958656311035) | ('z-direction', 0.7369154095649719) | ('DCE-MRI', 0.7509955167770386) | ('Musculoskeletal', 0.6045567989349365) |
| ('insufficiency', 0.6182920932769775) | ('xy-plane', 0.7321602702140808) | ('MRA', 0.7257292866706848) | ('osteoarticular', 0.5906753540039062) |
| ('ARF', 0.5905427932739258) | ('x-y', 0.7298304438591003) | ('diffusion-weighted', 0.7226229310035706) | ('dermatological', 0.5878671407699585) |
| ('heart', 0.5775583386421204) | ('axial', 0.7263980507850647) | ('CMR', 0.7127884030342102) | ('psychosomatic', 0.5870370864868164) |
| ('dysfunction', 0.561992883682251) | ('bisector', 0.7119565010070801) | ('Contrast-enhanced', 0.7117021083831787) | ('disabling', 0.5861454010009766) |
| ('failing', 0.561861515045166) | ('Z-axis', 0.7093360424041748) | ('resonance', 0.7098932862281799) | ('pain', 0.5821454524993896) |

## Averaged Cosine Similarity Scatter Plots




Using pre-trained **clinical embeddings** (above, (0.2-0.6)) instead of **one-hot encoding** (below, (0.85 - 0.975)) leads to lower similarity scores between category averaged answers which shows greater ability of clinical embeddings to the **capture semantics of answers.**




A similar trend is also found in averaged question embeddings although with a lower intensity due to common words such as ‘is’, ‘what’ etc. (refer to notebook for plots).

# Question Classifier : Cosine-Similarity of Clinical Embeddings

- For the total of 247 unique questions in the **training dataset**, clinical embedding vectors were calculated.
- Embedding of each question is the average embedding of all the words in it.
- For each question, the category was predicted by calculating cosine similarities with all other questions and **the category with maximum average cosine-similarity was predicted**.
- The model achieved an accuracy of **~79.75% i.e was correct 197 out of 247 times.**

# Next Up

1. Exploratory data analysis of images
2. Use the larger embedding pre-trained on PubMed + MIMIC notes
3. Train a baseline QA model : CNN, LSTM to encode image and text respectively.
4. Compare model on the different embeddings