統計的学習1　水野佳世子

Read the review https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5721660/ .

Is there any improvement in your skills to grab the content of the review?

If any, describe what they are and how they worked?

If no, describe how the lectures should be improved.

講義を受ける前では、論文中に出てくる専門用語のほとんどを理解できなかったが、講義を受けたことで、いくつかの統計的手法や概念は、何となくイメージを持って読み進めることができた。例えば、’support vector machines (SVMs)’ , ‘k-nearest neighbors (k-NN)’といった手法は、講義を受けていなかったら、全くわかっていなかったが、講義を思い出したり、講義資料を見直したりして、どのような手法かイメージしながら読むことができた。本論文で出てこなかった手法についても、講義資料を見直してさらに復習していこうと思った。’Truncated singular value decomposition (SVD)’, ‘probabilistic latent semantic analysis

(pLSA)’といった講義で出てこなかった用語は理解できなかったが、今後少しずつ勉強していきたい。

また、講義中に出てきた用語で、理解があやふやだったものも、論文を読んで、少し理解が深まったものもあった。例えば、’training set’, ‘test set’という用語について、講義でお聞きしたが、あまり具体的なイメージがつかめていなかったが、’training set’, `validation set’, ‘test set’は、subsetを分割して得られるもので、’training set’, `validation set’を使用してモデルをトレーニングし、最後に、モデルを‘test set’に当てはめて、その結果を評価する、という論文中の記載を読んで、具体的に考えられるようになった。

論文のTip9を読んで、今までの研究において、統計ソフトはSASを使用していたが、大学院卒業後は使用できなくなる可能性が高いため、講義を受講する際に、R, Pythonをダウンロードしたので、それを活かして、在学中にR, Pythonなどのフリーソフトも今後の自身の研究でも使用できるよう勉強していこうと思った。