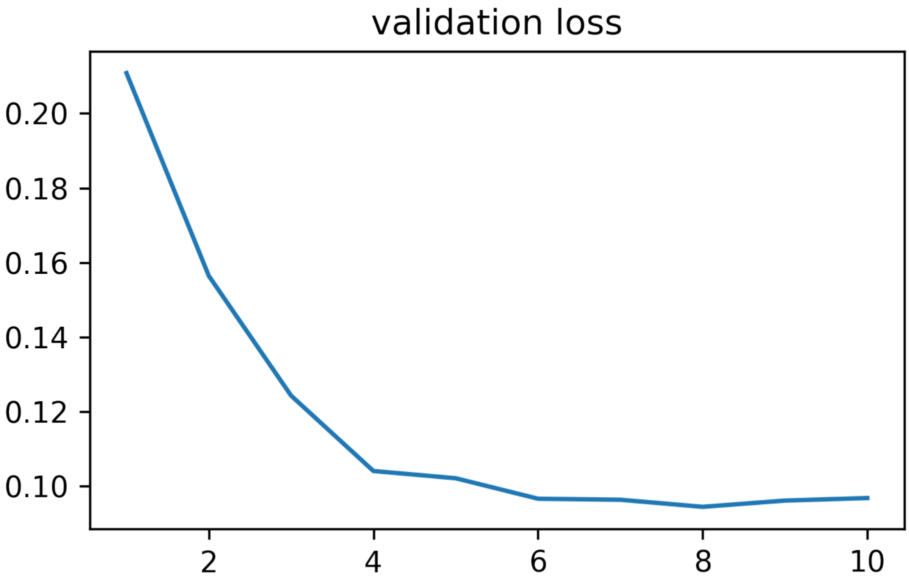
# 先端データ解析論　第 11 回小レポート

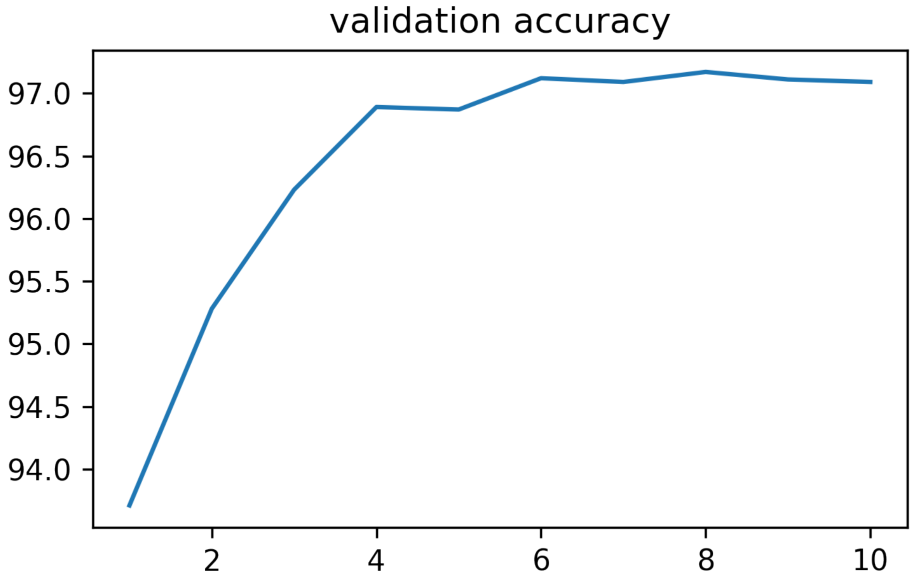
情報理工学系研究科電子情報学専攻 M1 堀 紡希 48216444

2021 年 6 月 22 日

## 宿題 1

図を以下に示す．optimizer として Adam を用いたためか, loss, accuracy ともに同じ epoch でも授業スライドで示された図よりも性能が良くなっている事がわかる．ソースコードを別の ipynb ファイルでも提出する．






## 宿題 2

多クラス分類問題において,

$$
\begin{aligned}
\delta_j^{(L)} &= \frac{\partial}{\partial u_j}\left(-\sum_{k=1}^{K} y_k \log f_k(\boldsymbol{x};\boldsymbol{w})\right) \\
&= -y_j \frac{\partial}{\partial u_j} \log f_j(\boldsymbol{x};\boldsymbol{w}) - \sum_{k\neq j} y_k \frac{\partial}{\partial u_k} \log f_k(\boldsymbol{x};\boldsymbol{\theta}) \\
&= -y_j \frac{\frac{\partial}{\partial u_j} f_j(\boldsymbol{x};\boldsymbol{w})}{f_j(\boldsymbol{x};\boldsymbol{w})} - \sum_{k\neq j} y_k \frac{\frac{\partial}{\partial u_j} f_k(\boldsymbol{x};\boldsymbol{w})}{f_k(\boldsymbol{x};\boldsymbol{w})}
\end{aligned}
$$

ここで,

$$
\begin{aligned}
\frac{\partial}{\partial u_j} f_j(\boldsymbol{x};\boldsymbol{w}) &= \frac{\partial}{\partial u_j} \frac{\exp(u_j)}{\sum_k \exp(u_k)} \\
&= \frac{\exp(u_j) \sum_{k\neq j} \exp(u_k)}{\left(\sum_k \exp(u_k)\right)^2}
\end{aligned}
$$

を用いると,

$$
\begin{aligned}
-y_j \frac{\frac{\partial}{\partial u_j} f_j(\boldsymbol{x};\boldsymbol{w})}{f_j(\boldsymbol{x};\boldsymbol{w})} &= -y_j \frac{\exp(u_j) \sum_{k\neq j} \exp(u_k)}{\left(\sum_k \exp(u_k)\right)^2} \frac{\sum_k \exp(u_k)}{\exp(u_j)} \\
&= -y_j \frac{\sum_{k\neq j} \exp(u_k)}{\sum_k \exp(u_k)} \\
&= -y_j \left(1 - \frac{\exp(u_j)}{\sum_k \exp(u_k)}\right) \\
&= -y_j(1 - f_j)
\end{aligned}
$$

また $k \neq j$ に対して,

$$
\begin{aligned}
\frac{\partial}{\partial u_j} f_k(\boldsymbol{x};\boldsymbol{w}) &= \frac{\partial}{\partial u_j} \frac{\exp(u_k)}{\sum_l \exp(u_l)} \\
&= -\frac{\exp(u_j)\exp(u_k)}{\left(\sum_l \exp(u_l)\right)^2}
\end{aligned}
$$

を用いると,

$$
\begin{aligned}
-\sum_{k\neq j} y_k \frac{\frac{\partial}{\partial u_j} f_k(\boldsymbol{x};\boldsymbol{w})}{f_k(\boldsymbol{x};\boldsymbol{w})} &= \sum_{k\neq j} y_k \frac{\exp(u_j)}{\sum_l \exp(u_l)} \\
&= \sum_{k\neq j} y_k f_j
\end{aligned}
$$

以上から,

$$
\begin{aligned}
\delta_j^{(L)} &= -y_j(1 - f_j) + \sum_{k\neq j} y_k f_j \\
&= \sum_k y_k f_j - y_j \\
&= f_j \sum_k y_k - y_j \\
&= f_j - y_k
\end{aligned}
$$