Learning Concepts using Deep Neural Networks

Sunil Kumar Vengalil,[1] and Neelam Sinha[2]
[1]*International Institute of Information Technology Bangalore*
[2]*International Institute of Information Technology Bangalore*

Abstract

The ML revolution is in full swing. In fact, the groundwork for it was prepared in the middle of the 20th century, yet, it is only with the ever continuing development of increasingly powerful computers, combined with computational algorithms refined over the past couple of decades, that the world has seen an explosion of applications of ML, in anything from health, to finance down to even autonomous cars!

Keywords: Machine Learning, Artificial Intelligence, Computer Science
DOI: 10.2018/JMLFS000001

1. INTRODUCTION

One of the major pain points of deep learning models, when used in industry, is the lack of explainability, i.e., they are unable to provide the exact reason for prediction. One the one hand, researchers build complex ensemble and deep architectures in order to increase the classification accuracy, which results in increased accuracy but the more complex the model is, the more difficult to provide explanations for the predictions. In many use cases, like medical, customer churn prediction etc, it is as equally important to get an explainable prediction as getting correct predictions. Sometimes one might even want to compromise with a less accurate model, if it can provide better explanations for the predictions. This behaviour of Machine Learning and Deep Learning models is in sharp contrast with how humans learn.

Learning in human posses the following special characteristics which the ML and DL models starve to achieve.

1. When humans learns something new, over a period of time they gradually 1)internalize the concepts 2) abstracts the concepts and 3) reuse the concepts in related tasks- all these accomplished usually from a very few samples [**?**].
2. The learned concepts are more generic as opposed the predictions, they make using the concept, in a particular scenario. Each scenario might be different, but the concept used to make an inference will be same. In other words, they use the same concept to make predictions for different input scenarios i.e the mapping between concept and scenario is one to many.
3. Humans can always provide an explanation for what they predict using the concepts they learned (Irrespective of the fact that the prediction might be wrong).

In this article, we examine how such behaviours can be incorporated into deep learning models. Explainability in DL models is something which has been actively researched and many approaches has been suggested in the literature in last decade[]. See section 2.1 for an extensive literature survey of explainability on deep learning models. However, our approach is not just to inject explainability into a trained deep learning model. We rather, focus on how the underlying model can learn **generalized concepts** which are used to make predictions.

There has been previous studies [**?**] on developing models that learn concept and use the learned concepts in a human like manner using Baysean Probabilistic Language [].

In this article we report results of experiemnts performed on image datasets of varying complexity starting from MNIST, CIFAR-10, CIFAR-100 and Cat Vs Dog datasets. We propose an approach for learning classification and segmentation tasks while the model also learns the concepts associated with the dataset/task.

2. RELATED WORK

2.1. Explainability in Deep Learning Models
2.2. Concept Learning
2.3. Semi-supervised Learning

3. PROPOSED METHOD

3.1. Overview of Approach
3.2. Datasets
3.3. Primary Visual Concepts
3.4. Composition of concepts

- `add tetex` allows access on `unity` to a comprehensive distribution of LaTeX called `tetex` (optional)

- Here, `ghostview` is used to view the final document
- Using instead `dvips -P pdf latex1` creates a postscript file that is optimal if the a pdf file is to be created, e.g., using acrobat distiller or the `ps2pdf` utility

```
stat% distill latex1.ps   OR  stat% ps2pdf latex1.ps
```

Structure of a `.tex` file:

- *Preamble*
  - Specify *document class* (article, report, book, letter, etc.)
  - Add any "*packages*" used (e.g., to import graphics, create headers and footers, etc.)
  - Specify *margins*, *indentation*, *spacing*, etc.
  - Define "*new commands*" (coming up…)
- *Document body*
  - The actual document content

Fun facts:

- % symbol is used to document the file or "*comment out*" text; anything to the right of a % does not appear in the document
- LaTeX commands start with \
- LaTeX is *case sensitive*

For example: Here is a sample preamble and document body for an article (See the web page for a full template file)

```
\documentclass[12pt]{article} % type size: also 10pt or 11pt
% commands to set margins and spacing -- all have defaults
\setlength{\textheight}{9in}    % height of text on a page
\setlength{\textwidth}{6.5in}   % width of text on a page
\setlength{\parskip}{2.3ex}     % space between paragraphs
% commands to invoke packages
\usepackage{graphicx,psfig,epsf}    % no limit to how many
% user-defined newcommands
\newcommand{\betahat}{\hat{\beta}}  % more on this shortly
%  start of document body

\begin{document}
\section{Introduction}              % sectioning command
This is the introduction...
\end{document}
```

Syntax: Some commands have arguments in braces { }, some do not

Some commands with no argument:

```
\ldots, \dag, \ddag, \%, \&, \#, \{ \}, \today, \LaTeX
```

…, †, ‡, %, &, #, { }, October 25, 2021, LaTeX

Commands with arguments: `\setlength{ ... }`, `\section{ ... }`, `\subsection{ ... }`, `\hspace{ ... }`, `\vspace{ ... }`

4. MODES AND ENVIRONMENTS

Modes: At any point in a LaTeX file, there is a current "*mode*" in effect

- *Paragraph mode* – the default text mode, with line wrap. A space between lines signals the start of a new paragraph
- *Math mode* – math symbols and commands may be used, and mathematical expressions result
- *LR mode* – "left-to-right" mode, lines do not automatically wrap around

Note on math mode: Math symbols and commands only work in math mode; if they are used in other modes, an *error* will result

Environments: Often, there is also an *environment* in effect that determines how material is displayed – the basic structure is

```
\begin{environment-name}
...
\end{environment-name}
```

For example: The `math` environment

```
the linear model
\begin{math}Y = X\beta + \epsilon\end{math}.
```

the linear model $Y = X\beta + \epsilon$.

- The popular shortcuts are to use `$ ... $` or `\( ... \)`, e.g.

```
the linear model $Y = X\beta + \epsilon$.
```

For example: Creating a numbered list

```
\begin{enumerate}
\item This is the first entry
\item This is the second entry
\item This is the third entry
\end{enumerate}
```

1. This is the first entry
2. This is the second entry
3. This is the third entry

Some popular environments:

Environment	Mode	Description
`math`	math	in-text mathematical expressions
`displaymath`	math	displayed mathematical expressions
`equation`	math	displayed expressions w/ line number
`eqnarray`	math	lines up equal signs, line numbers
`eqnarray*`	math	lines up equal signs, no line numbers
`array`	math	matrices and arrays
`itemize`	paragraph	list with bullets
`enumerate`	paragraph	list with numbers
`description`	paragraph	list with indentation
`tabular`	LR	align text in columns
`table`	paragraph	number and position table
`figure`	paragraph	number and position figure
`center`	paragraph	center text
`mbox`	LR	write text while in math mode

Math: LaTeX is *tailor-made* for writing involving high mathematical content! And it's easy!

- *Subscripts, superscripts, roots*

  ```
  e^y, x_{ij}, \sqrt{x+y}, \sum^n_{i=1}
  ```

  $e^y, x_{ij}, \sqrt{x+y}, \sum_{i=1}^{n}$

- *Greek*

  ```
  \alpha,\beta,\gamma,\delta,\epsilon,\eta,\theta,
  \lambda
  ```

  $\alpha, \beta, \gamma, \delta, \epsilon, \eta, \theta, \lambda$

  ```
  \Gamma,\Delta,\Theta,\Lambda,\Omega,\Sigma
  ```

  $\Gamma, \Delta, \Theta, \Lambda, \Omega, \Sigma$

- *Roofs*

  ```
  \hat{\alpha},\tilde{\alpha},\dot{x},\overline{x},
  \bar{x}
  ```

$\hat{\alpha}, \tilde{\alpha}, \dot{x}, \overline{x}, \bar{x}$

Math, continued:

- *Binary operations*

  `\pm,\times,\div,\cup,\otimes`

  $\pm, \times, \div, \cup, \otimes$

- *Relation symbols*

  `\leq,\subset,\in,\geq,\equiv,\sim,\approx,\neq,\perp`

  $\leq, \subset, \in, \geq, \equiv, \sim, \approx, \neq, \perp$

- *Arrows*

  `\rightarrow,\Leftarrow,\Leftrightarrow,\uparrow`

  $\rightarrow, \Leftarrow, \Leftrightarrow, \uparrow$

- *Miscellaneous*

  `\forall,\exists,\Re,\sum,\prod,\int`

  $\forall, \exists, \Re, \sum, \prod, \int$

Math, continued: `textstyle` vs. `displaystyle`

- Math *displayed* as equations may be carried out using the `displaymath`, `equation`, `eqnarray*`, `eqnarray` environments
- *Shortcuts* when equations are *not numbered*: `$$ ... $$` or `\[ ... \]`; e.g.,

  `$$\sum^n_{i=1} x_i^2 (Y_{ij}-z_i \beta)$$`

$$\sum^n_{i=1} x_i^2 (Y_{ij} - z_i \beta)$$

- Some symbols appear *differently* depending on whether they are in the text or displayed; e.g.,

  `$\sum^n_{i=1}$ VS. $$\sum^n_{i=1}$$`

  $\sum^n_{i=1}$ VS. $\displaystyle\sum^n_{i=1}$

- Can be *overridden* with `textstyle{ }` and `\displaystyle{ }`

Math, continued:

- *Products, integrals, unions*

```
$$\prod^n_{j=1},\hspace{0.1in} \int^\infty_t f(u) du,
\hspace{0.1in}\bigcup_{A: A \in \Omega}$$
```

$$\prod_{j=1}^{n}, \quad \int_t^\infty f(u)du, \quad \bigcup_{A:A\in\Omega}$$

- *Special functions*

```
$\exp(x), \log y, \sin(k\pi), \min_x f(x)$
```

$\exp(x), \log y, \sin(k\pi), \min_x f(x)$

- Fractions, partial derivatives

```
$$\frac{\exp(x^T \beta)}{1+\exp(x^T \beta)},
\frac{\partial u}{\partial x}$$
```

$$\frac{\exp(x^T\beta)}{1+\exp(x^T\beta)}, \frac{\partial u}{\partial x}$$

Note: Use `\displaystyle` for fractions; otherwise they are too small
Math, continued: There are different ways to present math in **boldface**; here are two

- `$\mbox{\boldmath $X$}$`, output $\mbox{\boldmath X}$
  `$\mbox{\boldmath $\Sigma$}$`, output $\mbox{\boldmath Σ}$
- `$\mathbf{X}$, $\mathbf{\Sigma}$`
  $\mathbf{X}$, $\mathbf{\Sigma}$

Math, continued: `array` and `eqnarray` environments

- (2×3) *matrix*:

```
\left( \begin{array}{ccc}
   x_{11} & x_{12} & x_{13}\\
   x_{21} & x_{22} & x_{23}
       \end{array} \right)
```

$$\left(\begin{array}{ccc} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{array} \right)$$

- *Determinant* of (2×2) matrix:

```
\left| \begin{array}{cc}
   a_{11} & a_{12} \\
   a_{21} & a_{22} \end{array} \right|
```

$$\left| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right|$$

Math, continued: `array` and `eqnarray` environments

- *Braces*

```
x = \left\{ \begin{array}{l}
     \sin x \mbox{ if } y<3, \\
     \cos x \mbox{ if } y \geq 3
            \end{array} \right.
```

$$x = \left\{ \begin{array}{l} \sin x \mbox{ if } y<3, \\ \cos x \mbox{ if } y \geq 3 \end{array} \right.$$

- *Binomial coefficients*: `\left( \begin{array}{c}N \\ y \end{array} \right)`

$$\left(\begin{array}{c} N \\ y \end{array} \right)$$

Math, continued: `array` and `eqnarray` environments

- *Equation with several lines, = signs lined up*

```
\begin{eqnarray*}
\Delta_i & = & \sum_j \sum_{k \neq j} \mbox{Corr}(Y_{ij},Y_{ik}) \\
 & = & \sum_j \sum_{k \neq j} \rho_i^{\parallel j-k \parallel} \\
 & = & \frac{2 \rho_i}{1-\rho_i} \left\{ n_i-1 -
          \frac{\rho_i(1-\rho_i^{n_i-1})}{1-\rho_i} \right\}
\end{eqnarray*}
```

$$\begin{eqnarray*}
\Delta_i & = & \sum_j \sum_{k \neq j} \mbox{Corr}(Y_{ij},Y_{ik}) \\
 & = & \sum_j \sum_{k \neq j} \rho_i^{\parallel j-k \parallel} \\
 & = & \frac{2 \rho_i}{1-\rho_i} \left\{ n_i-1 -
          \frac{\rho_i(1-\rho_i^{n_i-1})}{1-\rho_i} \right\}
\end{eqnarray*}$$

The `tabular` environment:

- As with `array`, separate *elements* with &, make *new line* with \\
- Specify *number of columns* and type of *justification* at top, add *vertical* and *horizontal* lines

```
\begin{tabular}{c|rr}
 & \multicolumn{2}{c}{Results} \\
Parameter & \multicolumn{1}{c}{Bias} & \multicolumn{1}{c}{SE} \\
\hline
$\beta_0$ & $-$0.030 & 0.12 \\
$\beta_1$ & 0.002 & 0.07
\end{tabular}
```

	Results	
Parameter	Bias	SE
β_0	-0.030	0.12
β_1	0.002	0.07

5. NEWCOMMANDS

Motivation: In technical typing, the same (nasty) expression may appear *frequently*

- A `newcommand` is like a "*shortcut*" to produce the expression easily
- `\newcommand{keyword}{text}`
- A `newcommand` declaration may appear *anywhere* in a LaTeX source file (preamble or body) and is defined thereafter
- A `newcommand` `keyword` may *not* contain numbers

Examples: Some `newcommand` definitions and their usage

```
\newcommand{\bbeta}{\mbox{\boldmath $\beta$}}
\newcommand{\betahatj}{\widehat{\bbeta}_j}
\newcommand{\var}{\mbox{var}}
\newcommand{\sumjn}{\sum^n_{j=1}}
```

- Note that a *previously-defined* `newcommand` may be used in defining a *new* `newcommand`

```
$$\sumjn \var(\betahatj)$$
```

$$\sumjn \var(\betahatj)$$

6. CROSS REFERENCES

Advantage: A *built-in* feature of LaTeX is that it *automatically* keeps track of sections, numbered equations, pages, and so on

- Sections, equations, tables, figures, pages etc. may be *labeled* and referred to by the label
- If new labeled entities are added, LaTeX *renumbers* them automatically
- It is even possible to generate a *table of contents* and *index* for a document
- To set up cross references correctly, must process a document *twice*

```
\LaTeX Warning: Label(s) may have changed.
   Rerun to get cross-references right.
```

Examples:

- Numbered equation

```
\begin{equation}
\var(\alpha) = \sumjn \var(\betahatj)
\label{eq:alpha}
\end{equation}

In equation~\ref{eq:alpha}, we see that...
```

Examples, continued:

- Section label

```
\section{Introduction}
\label{s:intro}

...As discussed in Section~\ref{s:intro},
kurtosis...
```

- Page label

```
Thus, we see that calculation of the variance is
straightforward \label{p:var}


...On page~\pageref{p:var}, the variance
 calculation...
```

7. PACKAGES

Useful utilities: LaTeX is much more *powerful* than the intrinsic features would suggest

- A *huge* user community
- Contributed *document classes*, "*add-ons*" to allow different capabilities and customization
- "*Packages*"
- Define new commands, syntax, etc.
- Visit `CTAN` (see slide 11)

Example: `fancyheadings.sty` – make "*fancy*" document *headers* and *footers*

- In preamble

```
\usepackage{fancyheadings}
\lhead{\footnotesize \bf CHAPTER \thesection}
\rhead{\footnotesize \bf ST 762, M. DAVIDIAN}
\cfoot{\footnotesize PAGE \rm\thepage}
```

- See `http://www.stat.ncsu.edu/∼st762_info/` for results

Example: `shadow.sty` – make "*shadowboxes*"

- In preamble

```
\usepackage{shadow}

\shabox{This stuff}
```

This stuff

In addition: There are also user-defined, alternative *document classes*

- *Journals, book publishers* may have their own class to create articles, pages with a specific format

Dissertations:
At NCSU, dissertations may be created in LaTeX using special a special style; to learn more, visit

```
http://www2.acs.ncsu.edu/grad/ETD/tutorial/latex.htm

http://www.stat.ncsu.edu/computing/howto/latex/
          session_2/session2.html
```

8. IMPORTING GRAPHICS

Numerous options:
We discuss three of these

- `psfig` – `\usepackage{psfig}`

```
\psfig{figure=dental.ps,height=2.5in}
```

- `epsf` – `\usepackage{epsf}`

```
\epsfysize=2.5in
\epsfbox{dental.ps}
```

- `graphicx` – `\usepackage{graphicx}`
- Can also import other formats (pdf, jpg, etc)

```
\includegraphics[height=2.5in]{dental.ps}
```

9. TABLES AND FIGURES

Two standard LaTeX environments:
`table` and `figure`

- Automatically *numbers* tables and figures
- Allow tables and figures to be formatted and *referenced* within a document
- Allow *captions*

```
\begin{table}[h!]
\tbl{Results of the simulation.\label{t:simresults}}{%
\begin{tabular}{crr}
\toprule
 & \multicolumn{2}{c}{Results} \\
 Parameter & \multicolumn{1}{c}{Bias} &
                              \multicolumn{1}{c}{SE} \\
\colrule
 $\beta_0$ & 0.030 & 0.12 \\
 $\beta_1$ & 0.002 & 0.07\\
 \botrule
\end{tabular}}
\end{table}
```

	Results	
Parameter	Bias	SE
β_0	0.030	0.12
β_1	0.002	0.07

TABLE 1: Results of the simulation.

- Reference – `In Table~\ref{t:simresults}, we see that...`
- In Table 1, we see that...

```
\begin{figure}
\centering
\includegraphics[height=2in]{fpo.eps}
\caption{The dental data of Pothoff and Roy.}
\label{f:dental}
\end{figure}
```

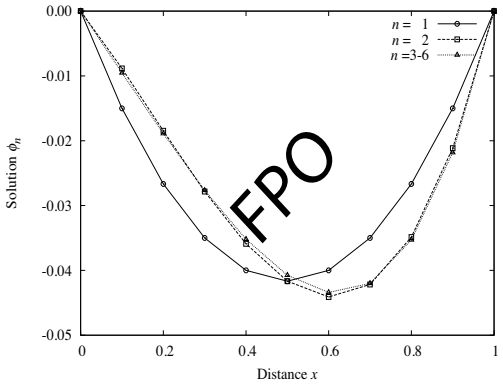


FIGURE 1: The dental data of Pothoff and Roy.

Useful package:

`subfigure` – `\usepackage{subfigure}`

- Create a "*multipanel*" figure from several files with each panel labeled

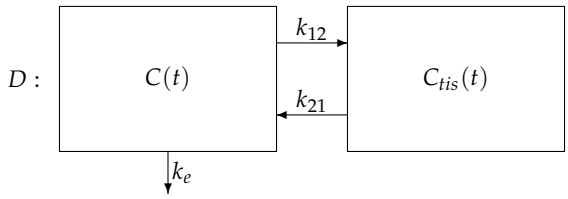
```
\begin{figure}
    \centering \subfigure[]{
        \includegraphics[width=1.5in]{dental.ps}}
        \hspace*{0.1in}
        \subfigure[]{
        \includegraphics[width=1.5in]{dental.ps}}
\caption{(a) The dental data of Pothoff and Roy.  (b) The dental
data of Pothoff and Roy, again.}
\label{f:dental2}
\end{figure}
```

10. PICTURES

LaTeX can "draw":

- `picture` environment
- The following is a *simple* picture – circles, curves, ovals, etc are also possible (see the documentation)

Two-compartment open model with IV administration:



$$\frac{dC(t)}{dt} = k_{21}C_{tis}(t) - k_{12}C(t) - k_eC(t),$$
$$\frac{C_{tis}(t)}{dt} = k_{12}C(t) - k_{21}C_{tis}(t),\ C_{tis}(0) = 0$$

Picture was made with:

```
\setlength{\unitlength}{1in}
\begin{picture}(5,1)
\put(0.5,0.5){\framebox(1.5,1){$C(t)$}}
\put(2,1.25){\vector(1,0){0.5}}
\put(2.25,1.35){\makebox(0,0){$k_{12}$}}
\put(2.5,0.75){\vector(-1,0){0.5}}
\put(2.25,0.85){\makebox(0,0){$k_{21}$}}
\put(2.5,0.5){\framebox(1.5,1){$C_{tis}(t)$}}
\put(0.25,1){\makebox(0,0){$D:$}}
\put(1.25,0.5){\vector(0,-1){0.3}}
\put(1.35,0.35){\makebox(0,0){$k_{e}$}}
\end{picture}
\end{center}
```

Other "drawing" resources:

- The `pstricks` package – really *intricate stuff* like grids, plots of functions, etc (see class web page for link to documentation)
- `xfig`

11. WHERE TO LEARN MORE

Books and guides:

- Lamport, L. (1994) *LaTeX: A Documentation Preparation System, User's Guide and Reference Manual* (The creator of LaTeX)
- Goossens, M. et al. (1994) *The LaTeX Companion*
- Kopka, H. (1999) *A Guide to LaTeX : Document Preparation for Beginners & Advanced Users*
- Hahn, J. (1993) *LaTeX for Everyone: A Reference Guide and Tutorial for Typesetting Documents Using a Computer*
- Oetiker, T. et al. (2002) *The Not So Short Introduction to LaTeX* 2_ε (*Available on the class web page*)

Resources online and on the Web:

- The *Comprehensive TEX Archive Network* (CTAN) `http://www.ctan.org` – a repository of tons of style files, packages, etc.
- Several *free* guides available on `unity` at `/afs/bp.ncsu.edu/ contrib/tetex107/share/texmf/doc/latex/general` (as `.dvi` or `.ps` files)
- Local intro tutorial `http://www.stat.ncsu.edu/computing/ howto/latex/session_1/`