Fast ai lesson 2 notes

23 June 2021 13:59

Lecture notes :

- If we did the training for the longer periods the model there are the chances that the model will overfit
- So to avoid that you have to train for smaller time frame and have both training and the validation data by which you can determine that the model is overfitting

- Learner :
  - the thing which tried to determine what type of the parameters are good for the training data
- The base model will be used to create the more complex model

Accuracy = 1 – error rate
Error rate is how much our model predicting the values wrong it get's printed after each epoch

* loss : the computer is using the measurement of the performance which computer is using for determining the values

- While the accuracy is the measurement where humans (we) are using to determine that the how model is performing

Loss measures every small bit of the information if he changed the parameter what will happen etc
but in turn the accuracy/ error rate matrix just measures how much predictions were right and how much predictions were wrong

Splitting of the validation data means our model never sees in training so it's completely untouched by it and it's not cheating any way

Book Notes :

The practice of the deep learning
There can't be just 6 lines of solution every time we code
We shouldn't overestimate and underestimate the capabilities of the deep learning
Cause if we underestimate them then we can't be able to solve the problems which are there really simple and can be solved with the great accuracy with deep learining
While on the other side if the problems are long and we assume that the deep learning will solve any of the problem then also there will be some hardship while solving the problem

- Try to build out projects even if you feel like the projects are not gonna get good cause you don't have dataset (good) remember the perfect doesn't exist you have to create your own dataset and then go with it

The deep learning is now used in :

- **Computer vision :**
  Deep learning is good at recognizing the patterns inside the images
  Even the data which can be converted into sound waves can be solved using the computer vision algorithms

- **NLP textual data :**
  The textual data can be also generated but since the data can be anything which can be generated textually we can have the scientific paper generated by the model but we can't guarrenty that the paper will be correct cause it's based on the trianed paper
  Another concern in NLP is there are models which can genearte the texts and which can recognize the computer generated text but it can be used for both of the models advantages (GAN) so that the model will in turn get better and the computer genearted text is almost same as the human genearted text

- **Combining the text and the images :**
  There text and the image model can be combined to get the subtitiles about the image informations but since as discussed previous we can't guarrently that the subtitles will be appropriate with the data *(image provided) cause the image data can be manipulated

  Eg the automatic system can be used to identify the stroke with help of CT scans

- **Tabular data**
  For analyzing the time series the tabular data is useful , dep learning is making the great strides . However it's generally used as a part of ensemble of multiple types of the model. If you already have a system that is using random forest of gradient boosting machines
  Then switching to the adding deep learning may not be result in the dramatic improvement on the other side the deep learning models take longer to train and predict than the simple machine learning models

- **Recommendation system :**
  The amazon is having great recommendation system but if you buy some autors book then it'll suggest you to buy that author's book more but it shouldn't be like that it should be focused on the category like the librarien of the library tells you that this writer also write some good fiction you can try reading his books

- **Another systems (another data types)**
  The systems such as the protein synthesis can be solved or identified similar to the natural languague processing models

Steps while building the model

- Define the objective for which you want to build the model
- Levers what the inputs for the model
- Where are you gonna collect the data for required inputs
- Models : create the actual model using the deep learning techniques

To turn the data into the dataloaders object we just have to know at least this four things

- What kind of the data is
- How to get the list of the items
- How to label this items
- How to create the validation set