CH 2 FROM MODEL TO PRODUCTION

Deep learning for Coders with fastai & PyTorch (source)

DL

CONSTRAINTS — CAPABILITES

underestimate ········· overestimate → poor results
overestimate ········· underestimate → didn't attempt a solvable problem
underestimate ········· underestimate → fail to consider and react to important issues
→ might not try things that may be beneficial

HOW TO START ?

1. select a project ▷ consider data availability !
2. start quickly ▷ don't wait for perfection in data, pipeline, etc.
3. create a full project pipeline ▷ iterate from end to end
   - ▷ if mobile app is the goal of the project ▷ have it after each iteration
   - ☆ see where the tricky bits are
   - ☆ gain better understanding of data needs; availability vs. needs
   - ☆ will have a working prototype to show
4. iterate in small increments ▷ document findings

STATE OF DEEP LEARNING (2020)

COMPUTER VISION

✓ recognize items in an image at least as well as people (even radiologists) ⇒ OBJECT RECOGNITION

✓ location of objects in an image; highlight the location and name each found object
⇒ OBJECT DETECTION
! image labeling can be slow and expensive
✓ synthetically generate variations of input images (e.g. rotating, ch. brightness, contrast, ...) ⇒ DATA AUGMENTATION

✓ categorizing every pixel ⇒ SEGMENTATION

✗ recognizing images structurally or in style diff. than the training images.
▷ OUT-OF-DOMAIN DATA ▷ learn how to manage for models in production

✓ convert non-image problem into a CV problem:
▷ classification of sound

NATURAL LANGUAGE PROC. (TEXT)

✓ classifing short / long documents ▶ spam Y/N, sentiment, author, website, ...

✓ generating context-appropriate text ▶ replies to social media, imitating author's style

✗ generating correct responses !

! used to spread disinformation ▶ create unrest ▶ encourage conflict

beware

(M1) text generation models always ahead of
(M2) models for recognizing automatically gener. text ▶ viscious circle (use M2 to improve M1)

✓ translate text / summarize long documents / find all mentions of a concept of interest ...

▶ protein chains as NLP problem !

TEXT & IMAGES

▷ train on images with captions ⇒ generate captions on new images ! always check whether the captions are correct

DL should not be used as an entirely automated process!

human interaction is essential

can make humans much more productive & more accurate

example DL system identifies potential stroke victims from CT scans ▷ send high-prority alert for the scans to be checked out by a human.

TABULAR DATA

- recently making great strides in DL
- ↑ variety of columns to include
  - ⇒ text / high cardinality categorical columns
- take long time to train vs. RF / GBM
  - ⇒ RAPIDS is changing this with GPU acceleration for the whole model pipeline

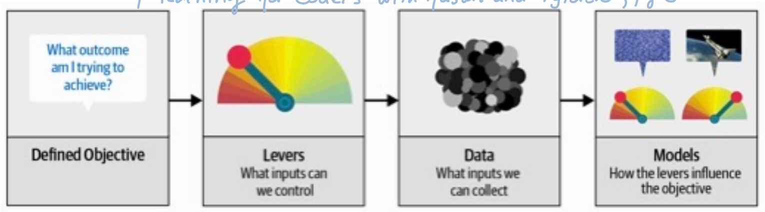
RECOMMENDATION SYSTEMS

- special type of tabular data → HCCV
  - high cardinality categ. var : users & products
- ✓ DL models are good at handling HCCV
- ! tells us which products a user might like vs. what recommendations would be helpful for a user.

THE DRIVE TRAIN APPROACH

⇒ consider how your model will be used in practice

Source: Deep learning for coders with fastai and PyTorch ; Pg 6



need a systematic design approach to build sophisticated data science products

! produce actionable results !

define a clear objective	what actions you need to take to meet your defined objective	what data you have; structure	build a model
ask what problem are you trying to solve			

Example: Google search engine

ask about what the user wants; ▷ the most relevant results; ▷ objective is to show them	rank the search resuts	what new data to consider; ▷ which pages linked to which other pages

Example: Recommendation System

drive additional sales by recommending items that would not be purchased otherwise	ranking of the recommendations	collect new data; ▷ conduct many randomized experiments (wide range of customers + recommendations	▷ 2 models for purchase probabilities conditional on seeing or not seeing a recommendation

difference of two models is the utility func. for a given recomm. for a customer.

GATHERING DATA

▷ find data online ▷ Bing Image Search | Duck Duck GO
▷ make sure your data is not biased
! consider what type of data your application will encounter
all data types should be included in the input data for the model

"**Actionable Auditing**: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI products"; Deb Raji

PROCESS

- download images
- verify images ▷ there are always some failed images ! ▷ unlink them
- structure data in a format suitable for training ▷ DataLoaders

DATA AUGMENTATION: random variation of input data
images appear different but do not change meaning of data
rotation // flipping // perspective warping // brightness changes // contrast changes } applied on images of the same size
batch_tfms

- train the model
- CONFUSION MATRIX ▷ check in which classes the model is making mistakes the most
▷ calculated using the validation set

		Predicted 0	Predicted 1
Actual	0	TN	FP
Actual	1	FN	TP

TN : true negative } correct predictions
TP : true positive }
FP : false positive : falsly predicting positive event
FN : false negative : falsy predicting negative event