With social interactions becoming more visual, social platforms must develop ways to detect and classify text, image and video content in order to surface relevant posts in search, identify trends and better detect potentially hateful content. On this front, Meta, Twitter and other social media companies have been working for years on its advanced image recognition, optical character recognition and natural language processing technologies.

Here's how we could develop a solution like this for social media contents.

Step 1 : Clarify requirements with the business team

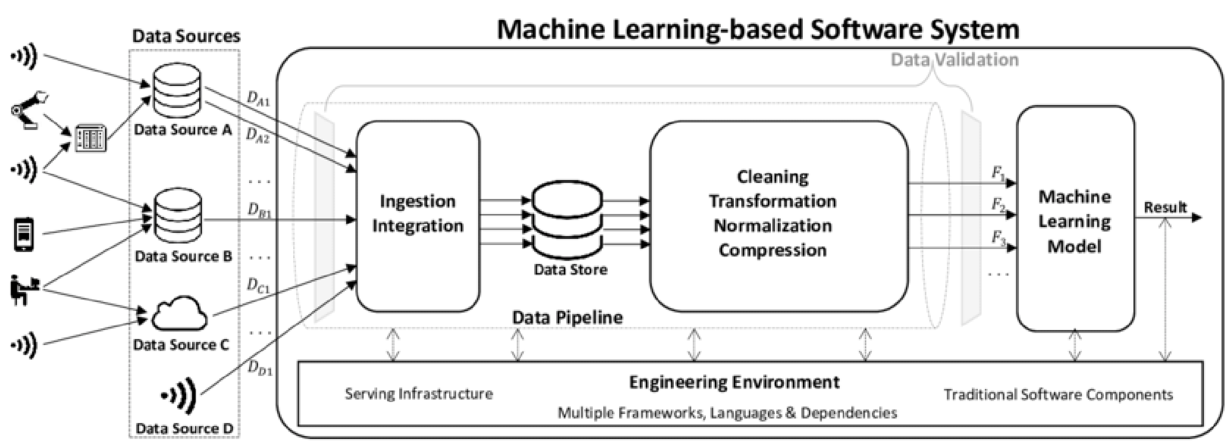
ML Production often involves getting engineers to think and prepare before coding to prevent losing time on things that won't function or solve the proper problem. Like traditional software engineering, machine learning is motivated by shiny technologies and novel concepts, but we may forget to step back to evaluate the actual problem we might solve, how users will interact with the product, and possible difficulties we can expect from the product.

Requirements Engineering encourages upfront thought about the problem, user needs, and potential issues. Requirements planning can prevent many costly issues later. After evaluating machine learning's suitability and defining the system, ML and non-ML components are separated. Similar to code review, requirements can be examined and audited.

In brief, requirements engineering can be divided into categories as follows:
1. When to use machine learning
2. Setting and measuring goal
3. Gathering requirements
4. Planning for mistakes

Step 2 : How ML model fits into the overall product backend

We need to think/draw a system diagram of input & output between backend and ml system.



D_{xi} ... Input Data Signal i from Data Source x   *F_i ... Feature i*

Step 3 : Data Related Activities

1. **Data Explore -** What's the dataset look like? In our case, as this is a social media agent, we need to store data in big data warehouse tools such as Redshift, BigQuery. Meta uses Hive. Exploratory analysis is performed by data scientists to better understand the data.
2. **Data Quality -** It is obviously important for machine learning as it influences the quality of the learned models. Must check data quality challenges such as : Accuracy, Completeness, Uniqueness, Consistency and Timeliness etc.
3. **Data Cleaning -** After data collection, data cleaning removes listed quality issues. For our case, our dataset will have multiple sources of warehousing. Multi-source difficulties can occur at both the schema (consistency, uniqueness) and instance (accuracy) levels. OpenRefine (previously Google Refine), Trifacta Wrangler are some commercial, academic, and open-source data cleaning solutions
4. **Data Schema -** ML data is commonly shared in schemaless formats including text entries in log files, tabular data in CSV files, tree-structured documents (json, XML), key-value pairs, etc. Data schema enforcement is crucial even for textual data sharing outside relational databases.
5. **Quality Assurance -** Finally, many data quality assurance procedures call for writing code in the system's machine learning pipelines and other parts. As always, it's critical to ensure that such code is accurate, reliable, and doesn't just crash on errors.
6. **Data Preparation -** Process data for ml model. Feature engineering is an important step to make the data compatible with ml models. Techniques such as encoding categorical variables, text embedding, converting image to tensor, data split into test train etc are some data preparation tasks.

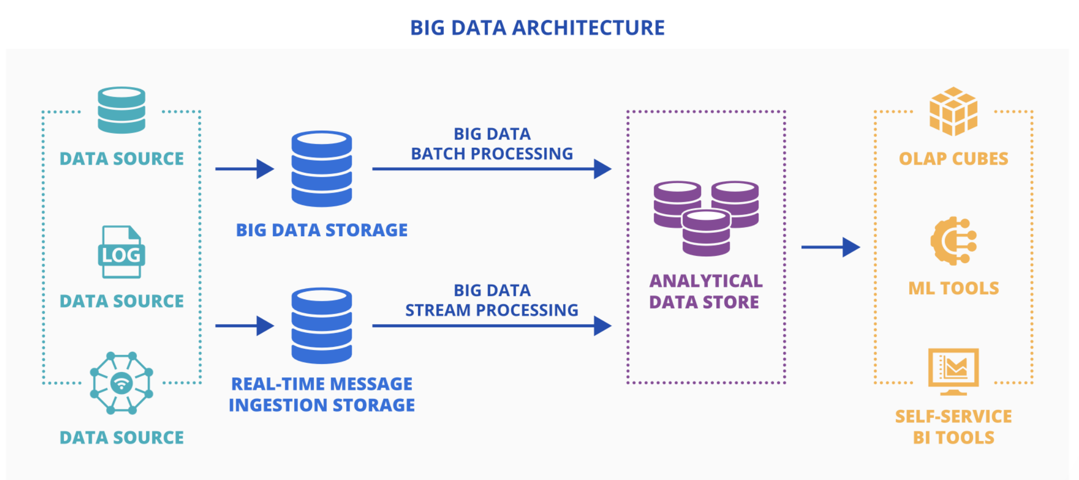


Fig : Typical data architecture for databases dealing with multi modality

Step 4 : Model Related Activities

1. **ML Pipeline -** As our model will take into account text, video and image data, we need to come up with a multi-modal approach for our agent. For text data, sequence models such as LSTM or transformer models such as BERT can be used for creating feature map. Also we need to improve the models with hyperparameter tuning and choosing the right layer structure. For the image model, object detection and image recognition algorithms can be used. For videos, Yolo , Detectron or other detection algorithms should be good to go. For other tabular data, we can come up with simple ml models such as XgBoost or simple NN structures. After that, we need to choose different hyperparameters in the models. Also we need to check bias-variance tradeoff which will give us an idea of overfitting vs underfitting.

2. **ML Model Debug -** While dealing with multiple input modalities like ours, we need to map each input into a feature space,concatenate and then train a neural network. Common bugs in ML development include incorrect shape of the tensors, incorrect preprocessing, toggling train/eval modes, controlling batch norm dependencies, numerical instability (NaNor or infs) etc.

3. **ML Model Deployment -** Steps involve rewriting the code (python to C# based on the requirements), enabling microservices such as RESTful APIs, Containerisation & Serialisation and loading the model in in memory key value store.

4. **Performance Monitoring -** A clearly defined performance monitoring system is necessary for every ML solution. An example of information that we might want to see for model serving applications includes: model identifier, deployment date/time, number of times the model has been served, predicted vs. actual/observed results etc

5. **Model Scaling -** Upon A/B testing with prior model (tradition or ML) and finalizing readouts with the authorities, we need to push for model scaling to more users and deeper neural complexity if required.

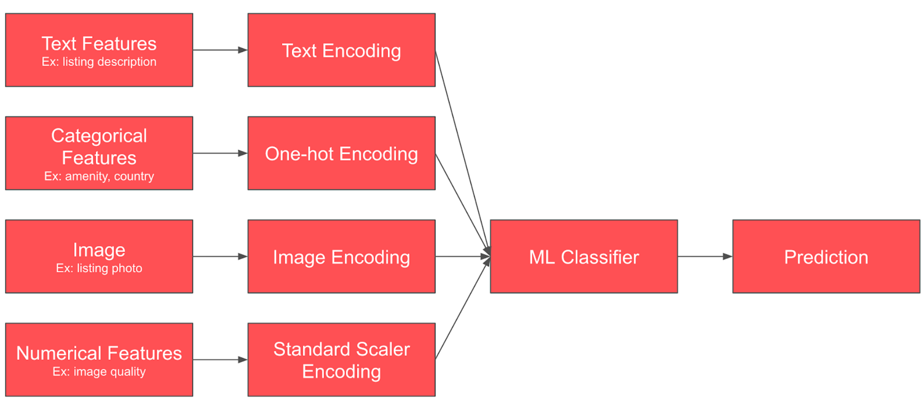


Figure : Multimodal model design for classification tye problems

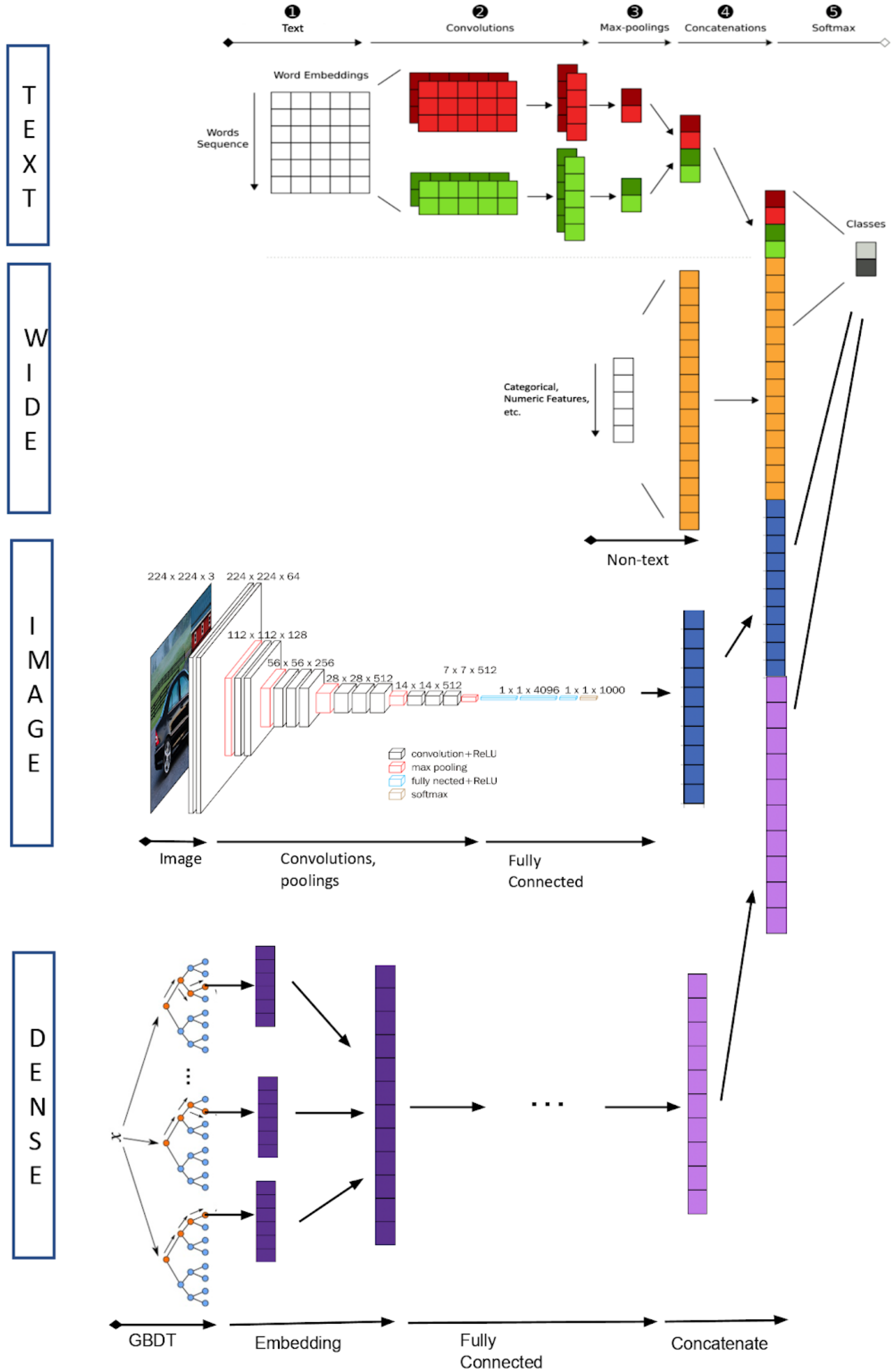


Figure : Visualization of the multimodal model having Text, Wide, Image and Dense channels Airbnb

References:

1. https://www.scnsoft.com/analytics/big-data/databases
2. https://medium.com/syncedreview/model-scaling-thats-both-accurate-and-fast-facebook-ai-proposes-novel-scaling-analysis-framework-249aec9bdfac
3. https://towardsdatascience.com/architecting-a-machine-learning-pipeline-a847f094d1c7
4. https://ckaestne.medium.com/machine-learning-in-production-book-overview-63be62393581
5. https://medium.com/@srank2000/how-facebook-handles-the-4-petabyte-of-data-generated-per-day-ab86877956f4
6. https://www.socialmediatoday.com/social-networks/facebooks-evolving-image-recognition-technology-offers-new-opportunities
7. https://www.socialmediatoday.com/news/facebook-can-now-understand-text-in-images-advancing-image-recognition-cap/532095/
8. https://github.com/chiphuyen/dmls-book
9. https://blog.hootsuite.com/social-media-data-collection/
10. https://medium.com/airbnb-engineering/widetext-a-multimodal-deep-learning-framework-31ce2565880c