

Transfer Learning
and
Applying Principles From This Class for LLM Use

26 March 2026

Alex Lyman

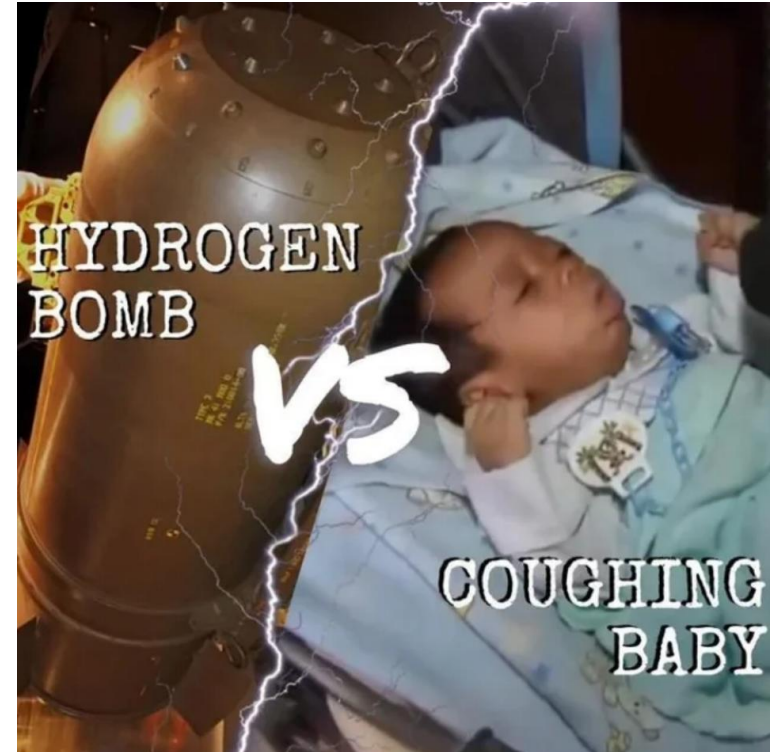
Transfer Learning

Transfer Learning (Pretrain and Fine Tune)

- Training a big ol' neural net from scratch takes a lot of work. (Lots of Compute)
- Big neural nets have lots of parameters.
- It takes lots of training examples to tweak all of those parameters into a useful direction. (Maybe even trillions)
- Why do that more than once?
- Solution: Pre-train a model on one general dataset, distribute for people to fine-tune.

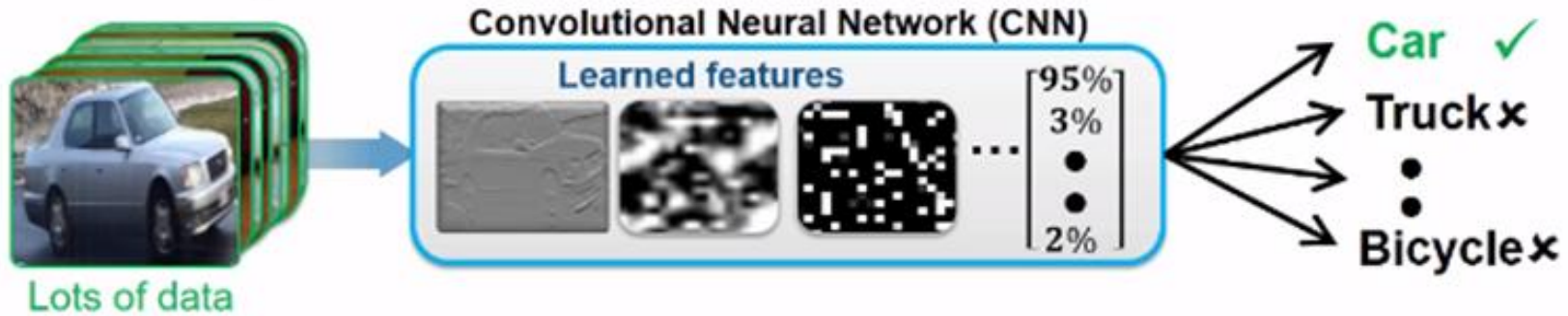
Transfer Learning (Pretrain and Fine Tune)

- Which of these is easier to teach college-level math to:
 - An 18-year old who's graduated from high school
 - A newborn baby
- Pretraining is like taking the model through high school
- Then fine-tuning is just a college class instead of raising a child to adulthood (training a model from scratch)



Approaches for Deep Learning

1. Train a Deep Neural Network from Scratch



Transfer Learning – Example 2



```
.run adven

WELCOME TO ADVENTURE!!  WOULD YOU LIKE INSTRUCTIONS?

yes

SOMEWHERE NEARBY IS COLOSSAL CAVE, WHERE OTHERS HAVE FOUND FORTUNES IN
TREASURE AND GOLD, THOUGH IT IS RUMORED THAT SOME WHO ENTER ARE NEVER
SEEN AGAIN.  MAGIC IS SAID TO WORK IN THE CAVE.  I WILL BE YOUR EYES
AND HANDS.  DIRECT ME WITH COMMANDS OF 1 OR 2 WORDS.  I SHOULD WARN
YOU THAT I LOOK AT ONLY THE FIRST FIVE LETTERS OF EACH WORD, SO YOU'LL
HAVE TO ENTER "NORTHEAST" AS "NE" TO DISTINGUISH IT FROM "NORTH".
(SHOULD YOU GET STUCK, TYPE "HELP" FOR SOME GENERAL HINTS.  FOR INFOR-
MATION ON HOW TO END YOUR ADVENTURE, ETC., TYPE "INFO".)
- - -
THIS PROGRAM WAS ORIGINALLY DEVELOPED BY WILLIE CROWTHER.  MOST OF THE
FEATURES OF THE CURRENT PROGRAM WERE ADDED BY DON WOODS (DON @ SU-AI).
CONTACT DON IF YOU HAVE ANY QUESTIONS, COMMENTS, ETC.

YOU ARE STANDING AT THE END OF A ROAD BEFORE A SMALL BRICK BUILDING.
AROUND YOU IS A FOREST.  A SMALL STREAM FLOWS OUT OF THE BUILDING AND
DOWN A GULLY.

east

YOU ARE INSIDE A BUILDING, A WELL HOUSE FOR A LARGE SPRING.


THERE ARE SOME KEYS ON THE GROUND HERE.






THERE IS A SHINY BRASS LAMP NEARBY.






THERE IS FOOD HERE.
```





Transfer Learning – Example 2

```
[GOAL] Welcome to Textworld! First, it would be fantastic if you could unlock the chest. And then, ensure that the chest inside the scullery is open. After that, take the Henderson's style passkey from the chest within the scullery. With the Henderson's style passkey, check that the Henderson's style locker is unlocked with the Henderson's style passkey. Then, ensure that the Henderson's style locker is open. Then, pick up the key from the Henderson's style locker. With the key, look and see that the gateway is unlocked with the key. After that, make it so that the gateway is open. After opening the gateway, make an effort to move west. With that done, ensure that the gate is open. That's it!  
[S] -- Scullery -- You've just shown up in a scullery. You see a locked Henderson's style locker here. You make out a locked chest. Oh wow! Is that what I think it is? It is! It's a refrigerator. You see a chair. The chair is normal. The chair appears to be empty. You move on, clearly furious with TextWorld. There is a closed portal leading south. There is a closed gateway leading west.  
[A] unlock chest with keycard  
[S] You unlock the chest.  
[A] open chest  
[S] You open the chest, revealing a Henderson's style passkey.  
[A] take Henderson's style passkey from chest  
[S] You take the Henderson's style passkey from the chest.  
[A] unlock Henderson's style locker with Henderson's style passkey  
[S] You unlock the Henderson's style locker.  
[A] open Henderson's style locker  
[S] You open the Henderson's style locker, revealing a key.  
[A] take key from Henderson's style locker  
[S] You take the key from the Henderson's style locker.  
[A] unlock gateway with key  
[S] You unlock gateway.  
[A] open gateway  
[S] You open gateway.  
[A] go west  
[S] -- Kitchen -- You're now in the kitchen. There is an open gateway leading east. There is a closed gate leading west.  
[A] open gate
```

 **Hugging Face**

 **Qwen / Qwen2.5-0.5B**   like 385  Follow  Qwen 76.1k

 Text Generation  Transformers  Safetensors  English  qwen2

 **Model card**  Files and versions  xet  Community **13**

- 97% on evaluation set of new text-based tasks
- On two-step tasks – First task 100%, second 70%

Getting LLMs to Do What You Want

Getting LLMs to Do What You Want

- You can use an LLM for pretty much anything.
 - Answering questions
 - Simulating people
 - Coding data
- How do you get the LLM to do what you want?
 - Fine-Tune (difficult, high control)
 - Prompting (easier, lower? control)
- How do you know the LLM is doing what you want it to do?
 - You should use a test set

Prompting With Train-Test Paradigm

Prompting With Train-Test Paradigm

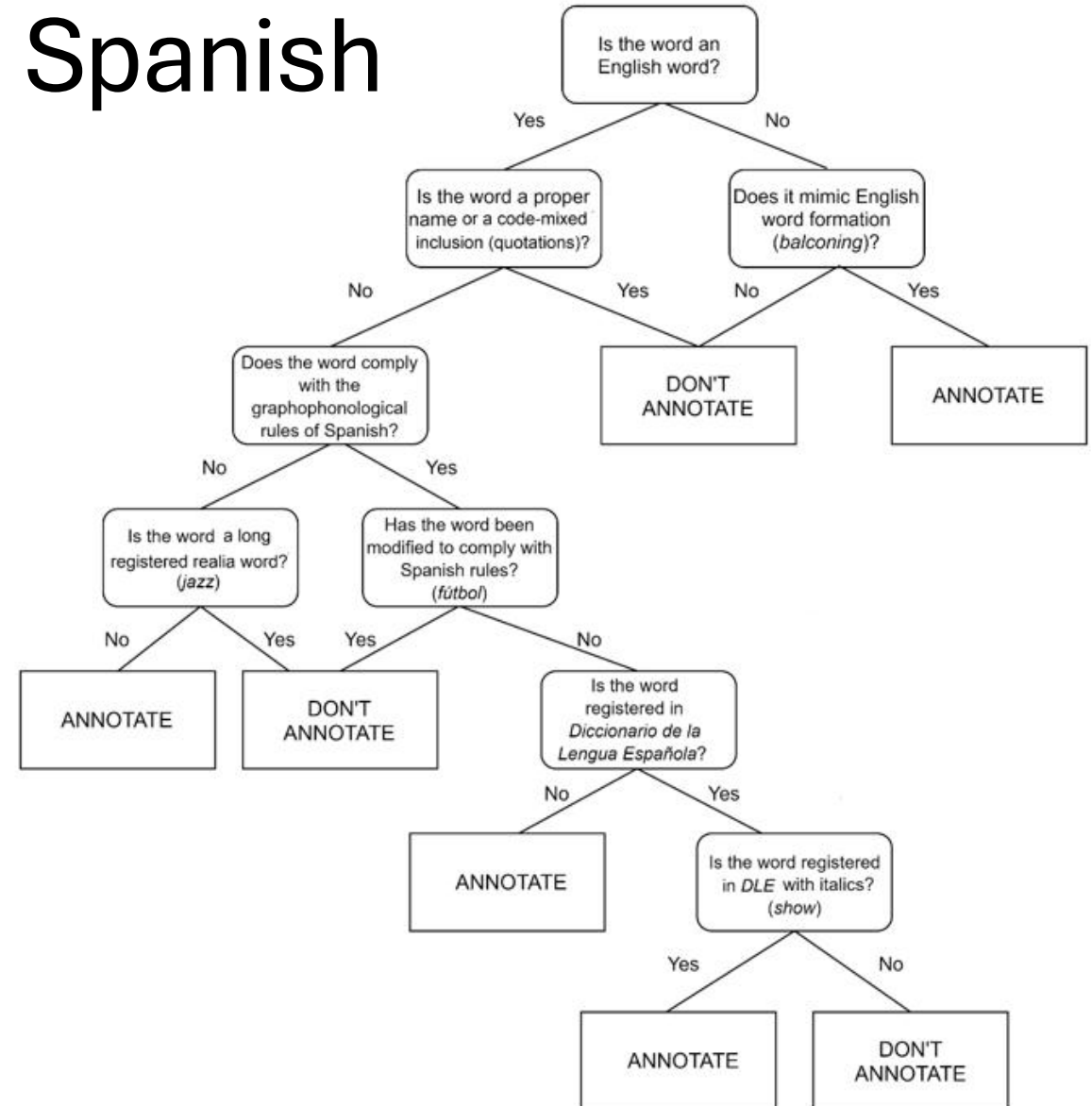
- Assume you want an LLM to do a task
- Example: Determine which Moral Foundations present in thousands of persuasive messages.
- How can you tell?
 - Hand annotate a small subset
 - Split into validation and test sets
 - While accuracy $<$ threshold:
 - Write a prompt
 - Test on validation set
 - Tweak prompt
 - Finally, test generalization on test set

- **Care/Harm**
- **Fairness/Cheating**
- **Loyalty/Betrayal**
- **Authority/Subversion**
- **Sanctity/Degradation**
- **Liberty/Oppression**

Case Study: Anglicism Detection in Spanish

Anglicism Detection in Spanish

- Anglicism – Word borrowed into Spanish that hasn't been grammatically incorporated into the language yet.
- Rules are really complex.
- How to detect?



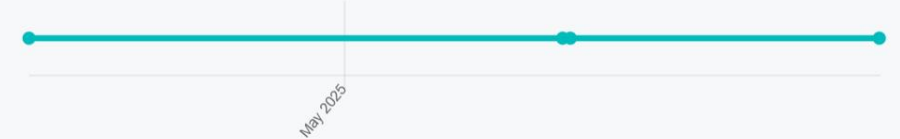
Anglicism Detection in Spanish

- Shared task at IberLEF 2025 (Iberian Language Evaluation Forum)
- Contestants were given a train and validation set with answers.
- Test set came without answers.
- I won (spoiler) with a .99 F1 score (beating a Chinese university team and a Latin America university team).



AUTOMATIC DETECTION OF BORROWINGS (ADOBO 2025)

ORGANIZED BY: [Jordip \(jordi.porta@uam.es\)](mailto:jordi.porta@uam.es)
CURRENT ACTIVE PHASE: None
CURRENT SERVER TIME: May 28, 2025 At 1:47 PM MDT
Docker image: [codalab/codalab-legacy:py37](#)



Task:	Results				ADoBo Final Task		
#	Participant	Entries	Date	ID	F1	precision	recall
1	qilex	1	2025-05-24 20:45	297754	0.99	0.99	0.99
2	shentzu	1	2025-05-24 12:38	297656	0.96	0.97	0.95
3	mheredia	1	2025-05-26 07:35	298600	0.93	0.93	0.94
4	HULAT-UC3M	1	2025-05-18 12:44	292548	0.92	0.96	0.88

Given the following Spanish sentence, identify all anglicisms.

An anglicism is a word or multi-word expression borrowed specifically from English that has recently been imported into the Spanish language and is used without orthographic adaptation.

Examples of anglicisms include 'running', 'smartwatch', 'influencer', 'country managers', 'marketing'.

This is the sentence to evaluate:

{sentence}

Output only the identified anglicisms, separated by semicolons. If no anglicisms are found, output 'None'.

The following rules provide a robust, step-by-step protocol for identifying emergent, unassimilated English lexical borrowings (anglicisms) in Spanish text.

1. Initial Identification

- Determine if the word or multiword expression is of English origin or mimics English word formation.
- If the word is not of English origin, do not consider it an anglicism.
- If it mimics English formation but does not exist in English (pseudoanglicism, e.g., balconing), consider it an anglicism.

2. Exclusion of Proper Names and Code-Mixed Inclusions

- If the word is a proper noun (person, organization, product, location, event, etc.) or a direct code-mixed quotation, do not consider it an anglicism.
- Borrowings embedded within proper nouns or named entities do not count, unless the proper noun is part of a multiword borrowing used grammatically as in English (e.g., Google cooking is annotated).

3. Graphophonological Compliance

- If the word's spelling and pronunciation conform to Spanish graphophonological rules (e.g., bar, club), proceed to dictionary checks.
- If not (e.g., show, look), generally consider it an anglicism unless it is a long-registered realia word (see Rule 5).

4. Adaptation and Assimilation Status

- If the word has been morphologically or orthographically adapted to Spanish (e.g., fútbol, tuit, lider), do not consider it an anglicism.
- If the word remains unadapted, continue to dictionary checks.

5. Dictionary Verification

- If the unadapted word is registered in the Diccionario de la Lengua Española (DLE):
- If it appears in italics, consider it an anglicism.
- If it appears without italics and with the relevant meaning, do not consider it an anglicism (it is considered assimilated).
- If it is not registered or not with the relevant meaning, consider it an anglicism.
- For multi-sense words (e.g., top), only consider them an anglicism when used with unregistered meanings.

6. Realia and Long-Registered Borrowings

- If the word is a long-registered realia borrowing (cultural terms like jazz, pizza, whisky, club), do not consider it an anglicism, even if unadapted.
- If the word is a recent or emergent realia borrowing not yet registered, consider it an anglicism.

7. Multiword Borrowings

- Do consider multiword expressions borrowed as a unit from English (e.g., reality show, best seller).
- For adjacent borrowings not forming a fixed English phrase (e.g., look sporty), select each word separately.

8. Exclusions and Special Cases

- Do not consider an anglicism:
- Latinisms, scientific units, species names, acronyms (unless part of a multiword borrowing), or digits in isolation.
- Metalinguistic usages, literal quotations, or code-switched expressions not integrated into the sentence.
- Names of peoples or languages, and words derived transparently from proper nouns (e.g., un iPhone, un whatsapp).
- Do consider an anglicism:
- Pseudoanglicisms (Spanish-coined words mimicking English, e.g., footing, balconing).
- Unadapted names of fictitious creatures (e.g., hobbit, troll).
- Borrowings embedded in compounds or prefixed forms, if the borrowed element retains independence (e.g., ex influencer, nano influencers).

The rules refer to dictionary checks. You don't have access to a dictionary, so do the best you can

Examples of words that are NOT anglicisms include:

- Words fully adapted to Spanish orthography (e.g., 'fútbol', 'líder', 'tuit')
- Long-established borrowings (e.g., 'bar', 'club', 'jazz', 'whisky')
- Proper names of people, places, or companies (e.g., 'Twitter' as a company name)
- Scientific terminology with Latin or Greek roots

Special rules for digital media and technology:

1. Names of online platforms (e.g., 'Facebook', 'Twitter') are NOT anglicisms when used as proper nouns to refer to the specific company or service.
2. However, these names ARE anglicisms when used generically (e.g., "Hizo un twitter" to mean "He made a tweet").
3. Generic terms related to digital culture ARE anglicisms when they maintain English form (e.g., 'post', 'blog', 'meme', 'podcast', 'streaming').
4. Terms for social media actions in their English form ARE anglicisms (e.g., 'like', 'share', 'tweet').

For multi-word expressions:

1. Analyze whether the entire expression functions as a unified borrowing from English.
2. Examples of multi-word anglicisms: 'fast food', 'big data', 'home office', 'fake news'.
3. Do NOT count individual English words that appear adjacent to each other but do not form a standard expression in English.
4. When in doubt about whether multiple words form a single anglicism, consider them separately.

For business, finance, and technical terminology:

1. Recent business terms that maintain English spelling ARE anglicisms (e.g., 'CEO', 'manager', 'marketing', 'branding', 'broker').
2. Technical computing terms that maintain English spelling ARE anglicisms (e.g., 'hardware', 'software', 'online', 'web').
3. Industry-specific English jargon ARE anglicisms (e.g., 'blockchain', 'big data', 'know-how').
4. International brand names and trademarks are NOT anglicisms unless used generically.

Follow this step-by-step process for identifying anglicisms:

1. First, identify all words and phrases that appear to have English origin.
2. For each candidate term, apply the exclusion criteria (proper names, adapted words, etc.).
3. For remaining terms, verify they maintain English spelling/pronunciation patterns.
4. Group related words that form a single expression in English.
5. Verify each term against the specific rules for digital media, business

Self-Refinement

- Have a different model see which kinds of mistakes the model was making on the validation set, and update the prompt with reminders based on errors.
- Have a smart reasoning model (OpenAI's o3) evaluate the examples.

Takeaways

- Principles of machine learning apply in new domains!
- There was no such thing as prompting an LLM when I learned about validation sets, but they are a good way to ensure LLMs are doing what you want/think they're doing.
- I don't know what that will look like for you, but the principles from this class will follow you through your ML career.