

Applied Machine Learning for Business Analytics

Lecture 12: Why do ML Projects Fail in Business

Logistics

- Thanks for your participation in the Kaggle Competition
- We have two quizzes for attendance check (2 points for each)
 - Even if you miss all, you will get 6 points for attendance
 - Scores would be released before next Friday
- Kaggle report would be due @ April 20th (**One more weekend**)
- Group project would be due @ April 27th
 - Naming of files!
 - Word and latex templates have been provided
 - https://bt5153msba.github.io/project/BT5153_ProjectGuidelines_Grading%20Criteria.pdf

Your Voice Matters

- Your Voice Matters
 - BT5153 requires training data for better performance in the future
 - “Student Feedback” under Account in Canvas
 - Thanks ahead for your time!

Agenda

1. Vibe Coding
2. Zillow Offers
3. Course Summary

1. Vibe Coding

Vibe Coding

- LLMs excels at writing, brainstorming and generating text
- It also highly effective at writing code
- With co-pilot features in IDE like cursor, there is a new concept popularized by Andrej Karpathy called **Vibe Coding**
 - It is a **prompt-driven** approach to software development where you write software by describing what you want in natural language, and let an AI generate the code



Andrej Karpathy
@karpathy



There's a new kind of coding I call "vibe coding", where you fully give in to the vibes, embrace exponentials, and forget that the code even exists. It's possible because the LLMs (e.g. Cursor Composer w Sonnet) are getting too good. Also I just talk to Composer with SuperWhisper so I barely even touch the keyboard. I ask for the dumbest things like "decrease the padding on the sidebar by half" because I'm too lazy to find it. I "Accept All" always, I don't read the diffs anymore. When I get error messages I just copy paste them in with no comment, usually that fixes it. The code grows beyond my usual comprehension, I'd have to really read through it for a while. Sometimes the LLMs can't fix a bug so I just work around it or ask for random changes until it goes away. It's not too bad for throwaway weekend projects, but still quite amusing. I'm building a project or webapp, but it's not really coding - I just see stuff, say stuff, run stuff, and copy paste stuff, and it mostly works.

7:17 am · 3 Feb 2025 · 4.4M Views



Generated by GPT-4o

Take Slack2JIRA Bot As an example

- Treasury team need to submit JIRA and write slack message for every crypto fund movement
 - It is happening every day (From Monday to Sunday)
 - However, the same content, just in different format.
 - Can we automate the flow from Slack to JIRA?

Saturday, March 8th

Hi @Rui

Please approve transfer from Omnibus to Cold Wallet as the ratio less than 70.5%. [JIRA](#)

1. 3500 SOL
2. 350 ETH
3. 400k XRP

Thank you.

PLUANTREAS-2833 FB rebalancing 8-Mar-25

Status: Backlog ☒ Type: Task ☐ Assignee: ZHAO RUI Priority: Normal

[Assign](#) [Change status](#) [Comment](#) [More actions...](#)

Added by Jira Cloud

1

FB rebalancing 8-Mar-25

[+ Add](#) [@ Apps](#)

Description

Hi @ZHAO RUI, please approve following transfer from Omnibus to Cold Wallet in [FB](#) as ratio less than 71%:

1. 3500 SOL
2. 350 ETH
3. 400k [XRP](#)

Use Vibe Coding

- Let us try to use vibe coding.
 - Yes, maybe some RPA softwares can handle this task
- After roughly 30 mins (vibe coding + joining in a meeting), POC was done
 - Never change or type any code base
 - Only talk with code-agent (via Cursor)
 - Give request
 - Refine request

Use Vibe Coding

```
Write an automation script to convert slack message into JIRA ticket.

Restore checkpoint

I'll help you create a script to convert Slack messages into JIRA tickets. First, let me create a Python script that will use both the Slack and JIRA APIs. I'll create a structured solution with proper error handling and configuration.

First, let's create a basic project structure with the necessary files:

$ mkdir -p src config tests

rz@Rays-MacBook-Pro JIRA_automation %

Now, I'll create a requirements.txt file with the necessary dependencies:

$ requirements.txt 1/1 ✓

slack-sdk==3.21.3
jira==3.5.1
python-dotenv==1.0.0
pyyaml==6.0.1

Let's create a configuration file to store our settings:

$ config.yaml 1/2 ✓
```

Initiate PRD



```
Make the name is really mentioned in JIRA ticket. Now, it is only plain text.

Restore checkpoint

I'll modify the code to create proper JIRA mentions using the [accountid] format. First, we need to get the JIRA account IDs for the mentioned users. Here's how we'll improve it:

$ slack_to_jira.py +53 -10 ✓

# Cache for user information
self.user_cache = {}
self.jira_user_cache = {}

def load_config(self):

Vice, let us fix the final issue: replace the reporter. How to solve the permission issue for JIRA API?

Restore checkpoint

I'll help you fix the reporter permission issue with the JIRA API. The error message indicates that the field 'reporter' cannot be set, which is typically a permissions problem.

Let's modify the approach to handle this properly:

$ slack_to_jira.py +12 -17 ✓
```

Refine PRD



```
Done, thank you

You're welcome! I'm glad we were able to successfully implement all the requested features for your Slack to JIRA integration script. Here's a summary of what we accomplished:

1. Fixed the reporter field issue by updating the approach to set it after ticket creation, with proper error handling and fallback mechanisms.
2. Implemented file attachment handling that successfully downloads files from Slack and attaches them to JIRA tickets.
3. Added support for custom summaries with the [summary] format at the beginning of messages.
4. Included date information in the ticket summaries using the DD-MM-YY format.
5. Set up a comprehensive logging system with:
   • Log rotation (10 MB max size, 5 backup files)
   • Detailed log information (timestamp, log level, process ID, hostname)
   • Both console and file output
6. Improved error handling throughout the script to ensure robustness.

The script is now ready for production use and should provide a reliable way to create JIRA tickets from Slack messages with proper attribution, file attachments, and detailed logging.

If you need any further adjustments or have questions about the implementation in the future, feel free to ask. Thank you for the opportunity to work on this project!
```

Retrospective Meeting

Use Vibe Coding



- Please approve following transfer in FB:
1. 205 USDT, from Omnibus to Refund Vault for user rewards distribution
 2. from User vault to Omnibus for Manual Sweep
 - a. 235.210266 USDC
 - b. 177.716109 USDC
 - c. 60.000001 USDC
 3. 0.011248353620806888 ETH_Base from Refund Vault to Sender's Address (require 1 approval) check

Slack message

Log

```
2025-03-18 17:34:20,818 - INFO - Message matches required format - will process
2025-03-18 17:34:20,818 - INFO - Found 1 valid messages that match criteria
2025-03-18 17:34:21,387 - INFO - Processing message from To Van Hung
2025-03-18 17:34:22,735 - INFO - Found summary in messages FB Approval test
2025-03-18 17:34:24,176 - INFO - Created JIRA ticket: PLUANTREAS-2921
2025-03-18 17:34:24,795 - INFO - Successfully updated reporter to To Van Hung for PLUANTREAS-2921
2025-03-18 17:34:25,025 - INFO - Added watcher ZHAO RUI to PLUANTREAS-2921
2025-03-18 17:34:25,338 - INFO - Added watcher Andiasa Adesia to PLUANTREAS-2921
2025-03-18 17:34:25,338 - INFO - Successfully created JIRA ticket PLUANTREAS-2921
2025-03-18 17:34:25,338 - INFO - Created JIRA ticket PLUANTREAS-2921 from To Van Hung's message in #test-jira
2025-03-18 17:34:25,339 - INFO - Script completed successfully at 2025-03-18 17:34:25
```

Projects / Pluang Treasury / Add parent / PLUANTREAS-2921

[18-03-25] Approval Request: FB Approval test

+ Add @ Apps

Description @ZHAO RUI

- Please approve following transfer in FB:
1. 205 USDT, from Omnibus to Refund Vault for user rewards distribution
 2. from User vault to Omnibus for Manual Sweep
 - a. 235.210266 USDC
 - b. 177.716109 USDC
 - c. 60.000001 USDC
 3. 0.011248353620806888 ETH_Base from Refund Vault to Sender's Address (require 1 approval) check

JIRA Ticket

Slack Message Details:

- From: @To Van Hung (to.vanhung@pluang.com)
- Channel: #test-jira
- Date: 2025-03-18 17:33:16

What might be the next

- More automation/software would be built
 - AI-assisted programming will enable non-tech ppls to write code & softwares
- There is still a gap between Vibe code and production-ready code
 - AI-generated code is often promising but incomplete
- As developers, use Vibe coding is actually more mentally exhausting
 - Since all of easy tasks have been automated, you will be responsible for the more challenging ones. Debugging AI-generated code would be more labor-intensive
 - All managers should have the same feeling :)
- It remains essential to understand the fundamental principles of software development
 - We have already seen more and more git repos or software projects unintentionally leaking API credentials.
 - You need to identify the source of bugs, even if you do not need to debug them by yourself.
- Maybe we can see more vibe marketing, vibe operation?
 - Execution would be automated.

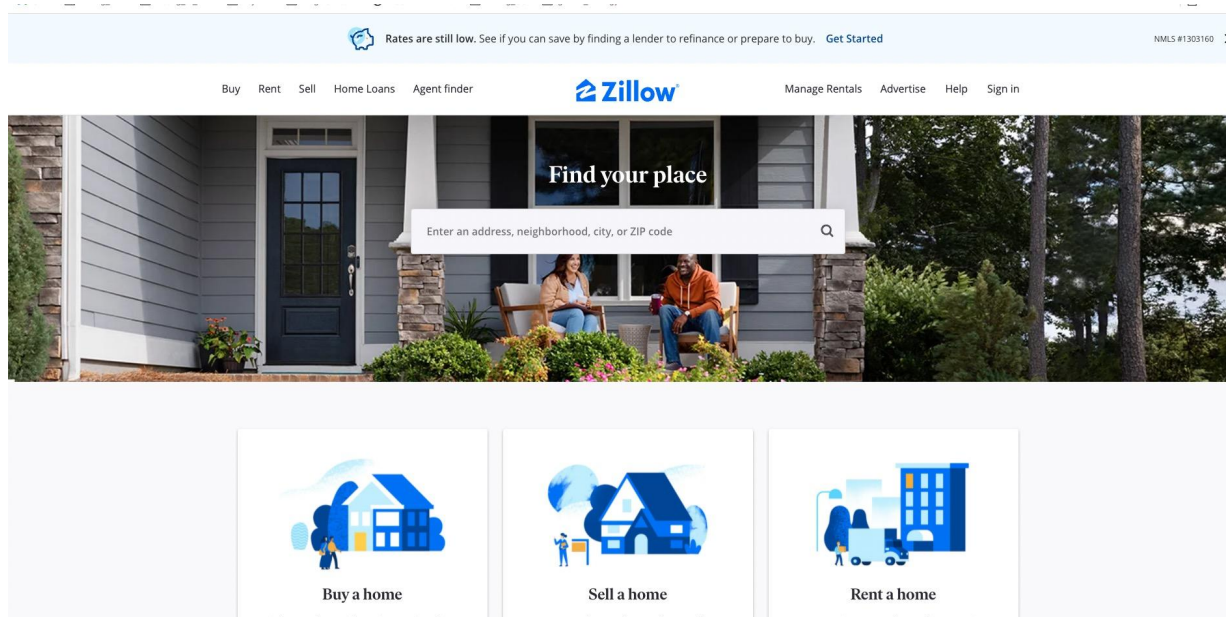
The Reality of Vibe Coding

<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2899	[14-03-25] Approval Request: FB Approval test	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2894	[14-03-25] Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2893	[14-03-25] Approval Request: FB Approval test	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2892	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2891	Approval Request: FB Approval test	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2890	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2889	Approval Request: ZHAO RUI-accountId:60744ebca5759fb0c069778678	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2888	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2887	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2886	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2885	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2884	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2883	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2882	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2881	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2880	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2879	Approval Request: following transfer in FB: 1, 205 USD; from Omnibus to Refund Vault for user rewards distribution 2. fro...	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2878	Approval Request: Approval Request	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	
<input type="checkbox"/>	<input checked="" type="checkbox"/>	PLUANTREAS-2877	Approval Request: this one	Unassigned	ZHAO RUI	Normal	BACKLOG	Unresolved	Mar 14, 2025	Mar 14, 2025	None	

2. Zillow Offers

Zillow

- an America online real-estate marketplace company



Zillow's Business Model



- How does Zillow make money?
 - <https://seoaves.com/zillow-business-model-how-does-zillow-make-money>
- Zillow Offers:
 - The system called “Zestimate” will analyze multiple data and predict the housing price as the bids for the seller
 - A home sale can be completed in a matter of hours
 - Due to the speed and convenience of the sale process, zillow can purchase houses below market value. After repairs and simple renovation, the houses could be sought to a new buyer at a higher price.
 - The delta in the price + commission fees charged from buyer and seller are the gross profits

1M Kaggle Competition

- At 2018, Zestimate launched a one-million kaggle competition
- In the competition, you will build machine learning models to predict the log error between the actual sale price and the Zestimate based on all the features of a home

$$\text{logerror} = \log(\text{Zestimate}) - \log(\text{SalePrice})$$

The screenshot shows the 'Zillow Prize: Zillow's Home Value Prediction (Zestimate)' competition page. At the top, it states the prize is \$1,200,000 and asks 'Can you improve the algorithm that changed the world of real estate?'. It mentions 3,770 teams and that the competition started 4 years ago. The page has tabs for Overview, Data, Code, Discussion, Leaderboard, and Rules. A 'Late Submission' button is visible. The 'Overview' section is active, showing a video player with a play button. Below the video, the address '111 Archer Ave, New York, NY 10031' is listed, along with '4 beds • 3 baths • 3,410 sqft'. The estimated price is '\$1,175,000' (Zestimate: \$1,279,448). The estimated mortgage is '\$4,461/mo'. A 'Get pre-qualified' button is present. The page also includes a map and contact information.

<https://www.kaggle.com/c/zillow-prize-1>





1M Kaggle Competition

- The winning solution pushed the Zestimate's current nationwide error rate of 4.5% to below 4%

Public Private

The private leaderboard is calculated with approximately 49% of the test data. This competition has completed. This leaderboard reflects the final standings.

■ Prize Winners

#	△	Team	Members	Score	Entries	Last	Code
1	▲ 9	Zensemble		 0.07408	253	4Y	
2	▲ 1	Juan Zhai 卷宅		 0.07421	53	4Y	

Meet the 'Zillow Prize' winners who get \$1M and bragging rights for beating the Zestimate

BY KURT SCHLOSSER on January 30, 2019 at 5:00 am

[Share](#) 102 [Tweet](#) [Share](#) [Reddit](#) [Email](#)



Zillow Chief Analytics Officer Stan Humphries, left, presents a check to Nima Shahbazi, a member of the team that won the Zillow Prize. (Zillow Photo)



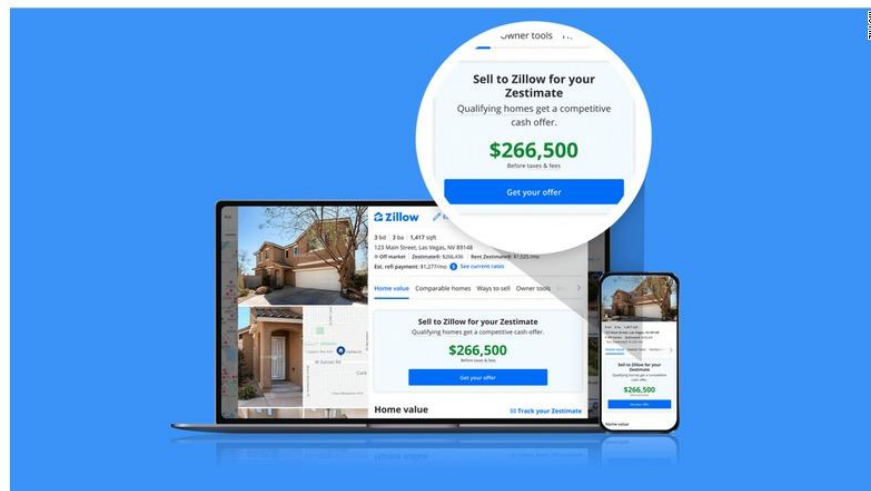
Zestimates was deployed to make offers

Zillow will now make cash offers for homes based on its 'Zestimates'



By Clare Duffy, CNN Business

Updated 1550 GMT (2350 HKT) February 25, 2021



Zillow's "Zestimates" will now represent initial cash offers to homeowners in some markets.

<https://edition.cnn.com/2021/02/25/tech/zillow-zestimate-cash-offer/index.html>

Stock Price from Oct 2021 to Mar 2022

Zillow Group Inc Class C ▲ **54.99** +2.00 (+3.77%)



What went wrong with Zillow Offers?

Zillow, facing big losses, quits flipping houses and will lay off a quarter of its staff.

The real estate website had been relying on its algorithm that estimates home values to buy and resell homes. That part of its business lost about \$420 million in three months.

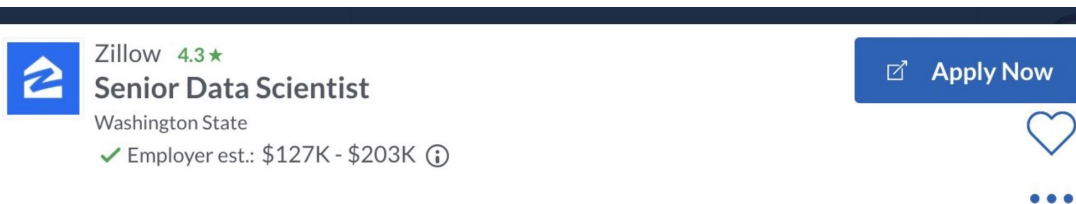
Zillow is sitting on thousands of houses worth less than what the company paid for them. Caitlin O'Hara for The New York Times

What went wrong with Zillow Offers?

1. Use ML to predict home prices
2. Use predicted prices to flip houses
3. ML models over-predict house prices
4. Buy houses at higher prices

Blaming game

1. Prophet: A python library for forecasting
2. Kaggle-style data science
3. Leadership
4. ML/DS team



A screenshot of a Zillow job listing. On the left is the Zillow logo (a blue square with a white 'Z'). To its right, the text reads 'Zillow 4.3 ★', 'Senior Data Scientist', 'Washington State', and '✓ Employer est.: \$127K - \$203K ⓘ'. On the far right is a blue button with a white arrow icon and the text 'Apply Now'. Below the button is a heart icon and three dots.

- Proven experience with Forecasting and Time Series modeling, especially Prophet, is strongly preferred.

More insightful discussion:

<https://ryxcommar.com/2021/11/06/zillow-prophet-time-series-and-prices/>

Prophet

- The model is developed by Facebook to predict the web traffic
 - Housing price do not have strong seasonality pattern which is different from web traffic

This is an accurate description of what Prophet's model is. To get a little more in the weeds, Prophet does the following linear decomposition:

- $g(t)$: Logistic or linear growth trend with optional linear splines (linear in the exponent for the logistic growth). The library calls the knots “change points.”
- $s(t)$: Sine and cosine (i.e. Fourier series) for seasonal terms.
- $h(t)$: Gaussian functions (bell curves) for holiday effects (instead of dummies, to make the effect smoother).

Adverse Selection

- Even the model accuracy is high, property owners will only sell when the predicted price is higher than their expected price



ML systems fail silently

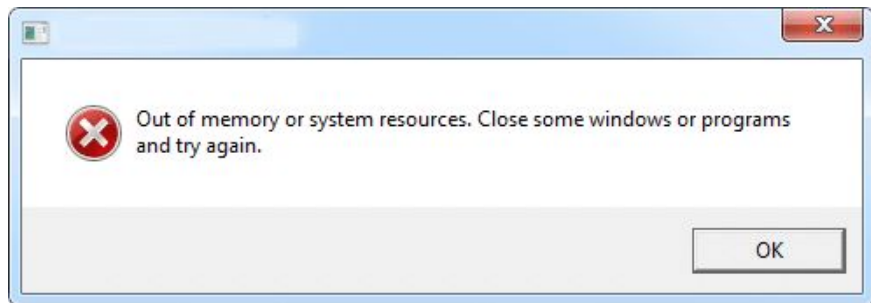
Normal softwares fail



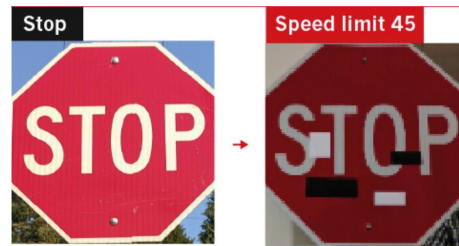
Stack Overflow is currently offline for maintenance

Routine maintenance usually takes less than an hour. If this turns into an extended outage, we will tweet updates from [@StackStatus](#) or post details on the [status blog](#).

```
^@^C^A>^D^A^@^P^@^C^AL^D^A^@^T^@^C^A^  
- stack_overflow^M  
^@^C^@R6003^M
```



ML systems fail



What is an ML failure?

A failure happens when one or more expectations of the system is violated.

Two types of expectations:

- Operational metrics: e.g. response time, downtime
- ML metrics: e.g. accuracy, MSE, BLUE score (machine translation)

What is an ML failure?

A failure happens when one or more expectations of the system is violated

- Traditional software: mostly operational metrics
- ML systems: operational + ML metrics
 - Ops: returns the risk scores of users within 800ms latency on average
 - ML: Accuracy as 80%

ML system failures

- If you call API to infer the user's risk score and get no response-> ops failure
- If the prediction is incorrect -> ML failure?

ML system failures

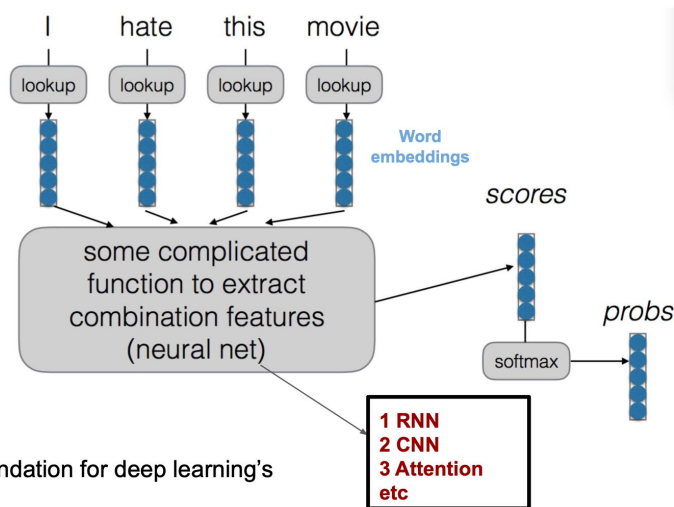
- If you call API to infer the user's risk score and get no response-> ops failure
- If the prediction is incorrect -> Might not be

ML failure when the predictions are **consistently** wrong

3 Course Summary


Recap: LLM

- From Bow to Word2Vec
- From Word2Vec to Transformers Auto-encoders
- LLMs



Word Embeddings is the foundation for deep learning's applications on NLP

Recap: Bridge the Gap

- Introduction to Machine Learning and its **Application**
 - Gap between theory and practice
 - **Data** Preparation
 - Machine Learning **Modelling**
 - Machine Learning **Evaluation**
 - Machine Learning **Deployment**
 - **Explainable** Machine Learning
 - Interpretability (remember the case: pokemon vs digimon or jpeg vs png)
 - Why do ML Projects **Fails** in Business
 - Business Understanding/Product Sense First
- 
- End2End ML System

Thank You

- Immense thanks to Xiaohui and Dingyu
- Enjoy having all of **you** in BT5153 this year. Appreciate your hard work!