

## 4NL3 Final Project Step 1

Neaha Bijo, William Clubine, Jasraj Johal, Raymond Ke

### Dataset Description

Our dataset is the Enron email dataset, a set of approximately 500,000 emails of employees of the Enron Corporation. It was obtained from Kaggle, with the dataset being found at <https://www.kaggle.com/datasets/wcukierski/enron-email-dataset/data>. It is an older dataset, last being updated in 2015. The dataset is in the form of a csv file with two columns, those being the file name (which is not relevant for our task), and the email content. The content contains the sender and recipients' emails, subject and message sent in the email, and some metadata such as the date the email was sent. For this step of the project, we selected 3000 emails at random from this dataset to annotate.

As a disclaimer, be advised that this dataset contains personal emails covering a wide variety of topics, including religion and politics.

### Annotation Guidelines

Your job as annotator is to determine which labels apply to each email in your given list of emails. The list of labels is as follows: Work, Personal, Transactions, Promotions, and Spam. Each email should be assigned either one or multiple of these labels.

**Work** - Work refers to emails directly related to company operations. This includes emails related to internal corporate communication, meeting scheduling, project discussions, reports, approvals, feedback, and corporate announcements.

Note that work emails may come from external parties (e.g., law firms, regulators, vendors) as long as the content is directly related to company business operations.

#### EXAMPLE:

Diana Willigerod 01/02/2001 08:48 AM To: Portland West Desk cc: Dan Dietrich/CAL/ECT
---

Subject: Network Update

As most of you know we have been experiencing some problems with our Network. Tonight starting at 7:30 Cisco doing some troubleshooting with our Networking team in Houston. I have co-ordinated the outages with the Real Time Schedulers.

Please feel free to contact me with any questions or concerns.

Thanks,  
Diana

This email is categorized as work because it relates directly to the company's operations without any transactional, promotional, personal, or spam messaging.

**Transactions** - Emails confirming purchases, orders, subscriptions, payments, receipts, invoices, shipping notifications, account changes, or any financial/commercial transaction.

**EXAMPLE:**

Message-ID: <9505762.1075855662432.JavaMail.evans@thyme>

Date: Fri, 11 May 2001 07:30:00 -0700 (PDT)

From: sdelgado\_vitoltvl@yahoo.com

To: peter.f.keavey@enron.com

Subject: 26JUL HOUSTON TO NEW YORK= TICKETED

Mime-Version: 1.0

Content-Type: text/plain; charset=us-ascii

Content-Transfer-Encoding: 7bit

X-From: sandra delgado <sdelgado\_vitoltvl@yahoo.com>

X-To: PETER.F.KEAVEY@ENRON.COM

X-cc:

X-bcc:

X-Folder: \Peter\_Keavey\_Jun2001\Notes Folders\Notes inbox

X-Origin: Keavey-P

X-FileName: pkeavey.nsf

AGENT JH/SS BOOKING REF

ZDJCNC

KEAVEY/PETER

ENRON  
1400 SMITH  
HOUSTON TX 77002  
PETE KEAVEY X37277

DATE: MAY 11 2001 ENRON

SERVICE DATE FROM TO DEPART  
ARRIVE

CONTINENTAL AIRLINES 26JUL HOUSTON TX NEW YORK NY 1058A 328P  
CO 1969 V THU G.BUSH INTERCO LA GUARDIA  
TERMINAL C TERMINAL M  
SNACK NON STOP  
RESERVATION CONFIRMED 3:30 DURATION  
AIRCRAFT: BOEING 737-800  
SEAT 06D NO SMOKING CONFIRMED KEAVEY/PETER

CONTINENTAL AIRLINES 29JUL NEW YORK NY HOUSTON TX 259P 558P  
CO 1975 V SUN LA GUARDIA G.BUSH INTERCO  
TERMINAL M TERMINAL C  
SNACK NON STOP  
RESERVATION CONFIRMED 3:59 DURATION  
AIRCRAFT: BOEING 737-300  
SEAT 10A NO SMOKING CONFIRMED KEAVEY/PETER

AIR FARE 241.86 TAX 26.64 TOTAL USD  
268.50

INVOICE TOTAL USD  
268.50

PAYMENT: CCCA5490850070001643/1103/A211782

RESERVATION NUMBER(S) CO/MVG8PL

KEAVEY/PETER TICKET:CO/ETKT 005 7028381411

CO FREQUENT FLYER COWH845429  
\*\*CONTINENTAL RECORD LOCATOR: MVG8PL  
THIS IS A TICKETLESS RESERVATION. PLEASE HAVE A  
ICTURE ID AVAILABLE AT THE AIRPORT. THANK YOU

\*\*\*\*\*  
NON-REFUNDABLE TKT MINIMUM \$100.00 CHANGE FEE  
THANK YOU FOR CALLING VITOL TRAVEL

-----  
Do You Yahoo!?

Yahoo! Auctions - buy the things you want at great prices

<http://auctions.yahoo.com/>

This email is categorized as a transaction because it is an e-ticket receipt confirming a completed airline ticket purchase (a transaction) with an invoice total (\$268.50), payment confirmation, and booking details.

**Promotions** - Marketing emails, advertisements, sales announcements, promotional offers, newsletters, product launches, discount notifications, or commercial solicitations.

**EXAMPLE:**

Message-ID: <17799625.1075844020039.JavaMail.evans@thyme>

Date: Thu, 6 Jul 2000 03:51:00 -0700 (PDT)

From: jay.patel@enron.com

Subject: Re: eSpeak Survey: The results are in!

Mime-Version: 1.0

Content-Type: text/plain; charset=us-ascii

Content-Transfer-Encoding: 7bit

X-From: Jay Patel

X-To: eThink@ENRON

X-cc: All Enron Worldwide@ENRON

X-bcc:

X-Folder: \Michelle\_Lokay\_Dec2000\_June2001\_1\Notes Folders\All documents

X-Origin: LOKAY-M

X-FileName: mlokay.nsf

No.10 on your list is Jack Welch, who is CEO of General Electric not General Motors. Please note this.

Regards

Jay N. Patel

EE & CC, Houston,TX

The eThink Team

From: eThink@ENRON on 07/05/2000 08:08 PM

Sent by: Enron Announcements@ENRON

To: All Enron Worldwide

cc:

Subject: eSpeak Survey: The results are in!

We recently asked you to create an eSpeak "wish list" for us. The response was tremendous and, in some cases, very creative. Your fellow employees asked for everyone from Britney Spears to George Bush. In all, we received requests for 61 persons outside of Enron to host an eSpeak event. We thought you might like to know who your colleagues find interesting, so we've provided a sampling of the suggestions below. Here are the top ten most requested eSpeakers.

10. Jack Welch, CEO, General Motors
9. Scott McNeally, CEO, Sun Microsystems
8. Satisfied Enron Customers
7. Stephen Covey, author, *The Seven Habits of Highly Effective People*
6. Oprah Winfrey, talkshow host
5. General Colin Powell, former Chairman, Joint Chiefs of Staff, U.S.A.
4. Alan Greenspan, Chairman, Federal Reserve, U.S.A.
3. Former U.S. President George Bush
2. Bill Gates, CEO, Microsoft
1. Texas Governor George W. Bush

While we can't make any promises, we will do our best to bring in at least some of these special guests. If you have any contacts who could help us get in touch with these guest speakers, please let us know at [ethink@enron.com](mailto:ethink@enron.com). Everybody needs a little help now and then, even eSpeak.

There were plenty of internal speakers requested, as well. Those requests should be somewhat easier to meet, although scheduling time with some of our colleagues can be quite a task!

Thank you all for your participation.

This email announces and promotes an upcoming company event and encourages participation from the employees. The results of the survey are shared to promote excitement about the event. The content of the email is not directly related to core business operations such as project tasks or deadlines, so it fits more in promotions than work. It is also not a one-to-one communication, so it does not fit under personal/social.

**Spam** - Unsolicited bulk emails, phishing attempts, suspicious offers, lottery/prize scams, get-rich-quick schemes, or emails with manipulative/deceptive intent.

**EXAMPLE:**

X-From: <jaym@igate.cbinet.com>  
X-To: dfoosum@enron.com  
X-cc:  
X-bcc:  
X-Folder: \Drew\_Fossum\_Dec2000\_June2001\_1\Notes Folders\All documents  
X-Origin: FOSSUM-D  
X-FileName: dfoosum.nsf

WorldEnergyOnline.com, is happy to provide you with a free trial subscription to Restructuring Today until December 12.

Restructuring Today has distinctive industry news and exclusive analysis -- and at only \$487 for over 250 issues each year, we consider it a bargain.

Act before December 15 to receive \$50 off!

Check it out, and if you'd like to subscribe, just hit "reply," and enter your billing info, or just call 781-939-2423.

You can also find a subscription coupon on the back page of the newsletter.

Requires Acrobat Reader 3.1 or higher, available free at:  
<http://www.adobe.com/products/acrobat/readstep2.html>

YES, I WANT TO SUBSCRIBE TO RESTRUCTURING TODAY IN PDF FORMAT FOR ONLY \$437/year. (\$50 off the regular price for subscription orders placed before December 15).

I WANT A CORPORATE SUBSCRIPTION TO RESTRUCTURING TODAY. The price is \$974 for the basic 5-reader site license, and \$121.75 for each additional reader. A universal site license for up to 100 readers is \$4000.

I want the basic site license plus \_\_\_\_\_ additional readers. Total = \$\_\_\_\_\_

I want a universal site license at \$4000 for up to 100 readers

NAME:

TITLE:

COMPANY:

BILLING ADDRESS:

PHONE:

EMAIL\*:

( ) PLEASE DO NOT SEND ME SAMPLES IN THE FUTURE

(\*For receiving Restructuring Today)

NOTE: This subscription offer is for first-time customers only -- current subscribers to Restructuring Today are not eligible to renew their subscription or order a site license through this offer.

Come visit us at <http://www.worldenergyonline.com>

[person\_id]  
- RT#1.pdf

This email is categorized as spam because it is unsolicited bulk promotional email marketing advertising a paid subscription service to someone who didn't request it, using time-limited discount tactics to pressure purchase.

**Personal** - Personal communication between individuals and unrelated to business. Includes emails related to personal relationships such as friends or family, social events, invitations, social media notifications, casual group communications, or social networking updates.

**EXAMPLE:**

Message-ID: <10740416.1075854426537.JavaMail.evans@thyme>

Date: Thu, 3 Aug 2000 01:36:00 -0700 (PDT)

From: darron.giron@enron.com

To: smmayers@earthlink.net

Subject: November

Mime-Version: 1.0

Content-Type: text/plain; charset=us-ascii

Content-Transfer-Encoding: 7bit

X-From: Darron C Giron

X-To: smmayers@earthlink.net

X-cc:

X-bcc:

X-Folder: \Darron\_Giron\_Jun2001\Notes Folders\sent mail

X-Origin: Giron-D

X-FileName: dgiron.nsf

Sean,

How's CO? I hope you are doing well. I know I have asked you before, but I seem to have accidentally deleted your dad and sister's email addresses. I want to ask you dad if he will be in town when KU plays Texas later in November. I want to bring the whole family. If you get a chance, shoot me a message with his info. My mom is vacationing in NY so I can't ask her.

Thanks.

DG

This is a personal/social email because it is entirely unrelated to work and discusses personal relationships (in this case, a coworker's dad and sister) and plans (watching KU in Texas).

**Multiple** - An email may receive up to 2 labels when it genuinely fits multiple categories. Assign the labels that best capture the primary purposes of the email. Do not over-label, a casual tone alone does not make an email Personal if the content is operational.

**Common multi-label combinations:**

- **Work + Transactions:** Internal memo forwarding a vendor invoice or receipt
- **Work + Personal:** Colleagues mixing business correspondence with social discussion in the same thread
- **Personal + Promotions:** A birthday party invitation sent broadly to an office group

**EXAMPLE:**

"Message-ID: <1598000.1075843425829.JavaMail.evans@thyme>

Date: Thu, 12 Apr 2001 11:57:00 -0700 (PDT)

From: jeff.dasovich@enron.com

To: cameron@perfect.com

Subject: Re: you & me & ian

Mime-Version: 1.0

Content-Type: text/plain; charset=ANSI\_X3.4-1968

Content-Transfer-Encoding: quoted-printable

X-From: Jeff Dasovich

X-To: Cameron Sellers <cameron@perfect.com>

X-cc:

X-bcc:

X-Folder: \Jeff\_Dasovich\_June2001\Notes Folders\All documents

X-Origin: DASOVICH-J

X-FileName: jdasovic.nsf

let's talk and figure out a time. i'm off for school now.

Best,

jeff

=09Cameron Sellers <cameron@perfect.com>

=0904/12/2001 05:44 PM

=09=09=20

=09=09 To: ""Jeff.Dasovich@enron.com"" <Jeff.Dasovich@enron.com>

=09=09 cc:=20

=09=09 Subject: you & me & ian

I would love for you to meet our new VP of Professional Services (who is, incidentally, now for all intents and purposes running our company). I think you will love him. He is very interested in the idea of working with Enron and I want you to meet him. Can we set up a dinner for the three of us sometime in the next couple weeks? I know your schedule is busy with school, so why don't you tell me a couple nights that work (other than Mondays) and I will make it work on our end. Please oh please.

Love,

Your Pal Cameron=20

P.S. Going to Colleen=01,s new (to me) place for some wine tonight. I=01,=

ll tell

her hello for you=01(

Cameron Sellers

Vice President, Business Development

PERFECT

1860 Embarcadero Road - Suite 210

Palo Alto, CA 94303

cameron@perfect.com

650.798.3366 (direct dial)

650.269.3366 (cell)

650.858.1095 (fax)"

Cameron is proposing a business dinner to introduce Jeff (Enron) to their new VP who "is very interested in working with Enron", a clear business development meeting (Work). At the same time, the email has a warm personal friendship tone ("Please oh please", "Your Pal Cameron") and closes with a purely personal social note about wine at Colleen's.

Note: Empty or unclassifiable emails: If an email has no body content and the subject alone is insufficient to classify, assign no labels (all zeros). This should be rare.

## Question 4

In total, each annotator annotated 870 instances, with 750 instances in the first round and an additional 120 instances in the second round. This took around 7 hours on average, or roughly 30.7 seconds per instance.

Through this process, a number of interesting data points were found. For example, we very quickly found instances that were blank or contained very little context to go off of, such as in the following example:

"Message-ID: <10143660.1075858856857.JavaMail.evans@thyme>

Date: Fri, 22 Jun 2001 12:16:02 -0700 (PDT)

From: stephanie.panus@enron.com

Subject: Angela Gregg

Mime-Version: 1.0

Content-Type: text/plain; charset=us-ascii

Content-Transfer-Encoding: 7bit

X-From: Panus, Stephanie </O=ENRON/OU=NA/CN=RECIPIENTS/CN=SPANUS>

X-To:

X-cc:

X-bcc:

X-Folder: \SPANUS (Non-Privileged)\Panus, Stephanie\Contacts

X-Origin: Panus-S  
X-FileName: SPANUS (Non-Privileged).pst

<Embedded Outlook Message Attachment>"

Upon discovering these, we regrouped and decided that such instances should be labelled with a zero in each category to mark that the email should be discarded in our actual training process later on. We discovered other cases that had context but were still somewhat ambiguous in terms of their category, like in the following example:

Message-ID: <16546169.1075856795235.JavaMail.evans@thyme>  
Date: Tue, 11 Jan 2000 00:55:00 -0800 (PST)  
From: robert.saltiel@enron.com  
To: vince.kaminski@enron.com  
Subject: CONGRATS!  
Mime-Version: 1.0  
Content-Type: text/plain; charset=us-ascii  
Content-Transfer-Encoding: 7bit  
X-From: Robert Saltiel  
X-To: Vince J Kaminski  
X-cc:  
X-bcc:  
X-Folder: \Vincent\_Kaminski\_Jun2001\_7\Notes Folders\Discussion threads  
X-Origin: Kaminski-V  
X-FileName: vkamins.nsf

Vince,

Congrats on your promotion! Well-done!

Rob

In this example, the email is somewhat work related (a promotion at work) but also somewhat personal (a personal congratulations for said promotion). In this case, the email could very well fit into either the work or the personal categories. This led to us developing our multiple annotation guidelines specifying when it would make sense to apply multiple labels. Another interesting example where this applies is shown below:

"Message-ID: <7711099.1075856075971.JavaMail.evans@thyme>

Date: Tue, 20 Feb 2001 06:52:00 -0800 (PST)

From: sally.beck@enron.com

To: cindy.olson@enron.com

Subject: Re: NABA photo shoot

Mime-Version: 1.0

Content-Type: text/plain; charset=us-ascii

Content-Transfer-Encoding: 7bit

X-From: Sally Beck

X-To: Cindy Olson

X-cc:

X-bcc:

X-Folder: \Sally\_Beck\_Jun2001\Notes Folders\Sent

X-Origin: Beck-S

X-FileName: sbeck.nsf

I am looking forward to involvement with the NABA and the prospect of hiring some strong candidates. If you want me for the photo shoot, February 23 looks good. I will be out of the office on March 2nd - two days of vacation on March 2 and 5 to chaperone Meagan's drill team to state competition in Dallas. With 90 girls on the team, that should be fun but exhausting!

It was great to see you and Grady last night at the rodeo. Let's definitely plan on a ball game in April as a chance to catch up. I just checked the calendar, and it looks like we are on for lunch on March 28!! A few weeks away, but at least it is on the calendar! If your schedule were to free up suddenly, just give me a call and I can see if I could change mine to have lunch sooner. --Sally

Cindy Olson@ENRON

02/19/2001 06:47 AM

To: Sally Beck/HOU/ECT@ECT

cc:

Subject: NABA photo shoot

Sally thanks so much.....I believe this is a great way to recruit talented African Americans.....I appreciate your sponsorship!!!!!!

----- Forwarded by Cindy Olson/Corp/Enron on 02/19/2001

06:36 AM -----

Gwynn Gorsuch

02/16/2001 04:50 PM

To: Cindy Olson/Corp/Enron@ENRON

cc: sally.beck@enron.com, wes.colwell@enron.com, Ted C Bland/HOU/ECT@ECT, Meliza Evanson/HOU/EES@EES, Zakiyyah McClure/HOU/EES@EES, Kim Cadena/HOU/EES@EES, Laura Schwartz/Corp/Enron@Enron, Karen Moore/NA/Enron@ENRON, Debbie Nowak/HR/Corp/Enron@ENRON, Gerry Gibson/Corp/Enron@Enron, Donald Miller/Corp/Enron@Enron

Subject: NABA photo shoot

EES and Sally Beck's organization in EWS are joining forces to underwrite Enron's recruiting presence at the 30th Annual National Association of Black Accountants National Conference and Job Fair this August in Dallas, Texas. NABA is an professional organization of 100,000 African Americans in the fields of accounting, finance, consulting and IT.

We are planning a photo shoot of a check presentation to the Executive Director of NABA, who is making a special trip to Houston for this purpose. The photo and caption will appear in the pre-conference NABA newsletter to its members (not a paid advertisement). We would like to have Cindy, Sally, Wes, Ted and the representative from EES in the photo, if possible. We also need to decide whose signature should appear on the mocked up check for the photo - what do you think?

We are checking on dates for the shoot now, and Debbie Nowak will be contacting you/your assistants to check for your availability. (likely contenders are Feb 23 around mid-day and March 2)

We are excited about Enron's participation in this conference and look forward to some great hires.

Thanks,

Gwynn

Here, the email thread begins primarily work related (discussing Enron's participation in a conference) but in later replies becomes more personal (talking about personal meetups). Based on this, we modified the multiple label guidelines to explicitly state that such cases should receive both a work and personal label.

By following the annotation guidelines carefully and updating them as needed, our team was hopeful that we would achieve relatively high levels of agreement. We used Krippendorff's Alpha to measure this on a per category basis and obtained the following results:

Category	Work	Personal	Transactions	Promotions	Spam
Krippendorff's Alpha	0.757	0.786	0.467	0.647	0.624

This shows a moderate to strong level of agreement in most categories, with especially high agreement in the work and personal categories. The other three categories were slightly more ambiguous and thus saw slightly lower agreement, with the transactions category being a notable outlier. On further inspection, however, we found that there were only 13 instances of disagreement in the transactions category. The low alpha comes mostly from the fact that there were not many transactions emails overall, with only 25 positive labels across both annotation rounds, resulting in an extremely low value for expected disagreement. Overall, we were happy with these results and feel confident in the quality of our annotations.

## Question 5a

Through the annotation process, We realized that email classification is much more complex than it initially appears. At first, it felt like a straightforward task, just read the email and assign one of the predefined labels. However, once we started annotating, we noticed that many emails contained overlapping signals. For example, some messages used business language while also referencing payments, and others had a promotional tone but were actually legitimate internal announcements. This made it clear that keywords alone are not enough, understanding the communicative intent behind the message is much more important.

We also learned that real-world data is rarely cleanly separable. Even with only five labels, we encountered several boundary cases where the distinction between categories was not immediately obvious. This exercise showed us how important it is to define precise and operational criteria when designing an annotation scheme. Over time, we shifted from relying on intuition to following a more structured, rule-based reasoning process. That shift helped us better understand what each label truly represents and improved consistency in our annotations.

## Question 5b

We expect models to struggle with several aspects of this dataset. One major challenge is label overlap. Some emails genuinely serve multiple purposes, such as combining a payment request with event participation details. This means the task is not purely single-label classification, and a model must handle multi-label outputs effectively.

Another difficulty is semantic ambiguity. Certain words like “offer,” “invoice,” or “announcement” appear across multiple categories. For example, “offer” could indicate a legitimate promotion or a suspicious spam message, depending on context. A model relying too heavily on surface-level keywords may misclassify these cases.

Class imbalance is also a concern. Since Work emails make up a large portion of the dataset, a model may become biased toward predicting Work more frequently, even when another label would be more appropriate. Finally, distinguishing between Promotions and Spam requires understanding tone, credibility and intent, something that goes beyond vocabulary alone. Subtle boundary cases that humans resolve using contextual judgment may be particularly difficult for models, especially when the meaning depends more on interpretation than on obvious lexical cues.

## Question 5c

While we always expected a need for multiple labels, we had mostly anticipated combinations of work plus transactions, promotions, or spam and combinations of personal plus transactions, promotions, or spam, so it was somewhat surprising when we found a large number of emails for which both the work and personal categories applied (see examples in Question 4). We were also surprised that there were relatively few transactional or spam emails, though in hindsight it makes sense that these would be drowned out by more frequent work threads. Overall, the actual content of these emails was not too surprising (except for the occasional humorous personal email) and was mostly focused on the operations of Enron.

## Question 5d

Based on our annotation experience, we expect several features to be particularly useful for classifying these emails. The email body text is the most important feature, as communicative intent is almost always expressed there rather than in metadata. Within the body, n-grams and key phrases tied to specific categories will carry strong signals, such as “invoice total,” “confirmation number,” and “wire transfer” for Transactions, or “act fast,” “limited time,” and

"you have won" for Spam. The sender and recipient domain (e.g., whether both parties are @enron.com versus an external marketing domain) provides a useful contextual signal for distinguishing Work from Personal and identifying likely Spam or Promotions senders. The subject line is a compact summary of intent and often resolves ambiguity when the body is short or missing. Distribution list size inferred from the To and Cc fields can help separate Personal (narrow, individual) from Promotions (broad, company-wide). Finally, email thread context such as forwarded or reply chains may be useful, since as we observed, a thread can begin as a Work exchange and shift into Personal discussion, making positional signals within the thread relevant for multi-label cases.

## Question 5e

Over time, we developed a consistent mental decision process that helped make the annotation more systematic. We would first identify the primary communicative purpose of the email. Then, we would ask whether the message involved active financial processing, if so, it would be labeled "Transactions". If it was not directly financial, we would determine whether it was related to professional or operational matters, which would fall under "Work". If the email was promoting a product or service, it would be labeled "Promotions", whereas messages that appeared deceptive, suspicious, or fraudulent would be categorized as "Spam". If the email originated from a platform and was about social engagement or notifications, it would be labeled "Social". Finally, if none of these applied and the communication was purely interpersonal, it would be classified as "Personal".

Instead of relying solely on surface-level keywords, we focused on identifying the underlying intent and functional role of each message. Developing this structured decision framework reduced ambiguity in borderline cases and improved consistency across annotations.

## Question 5f

One area that required some careful reading was distinguishing between the assignment guideline instructions and the annotation instructions that we were responsible for creating ourselves. Since both are referred to as "instructions," it initially took a moment to recognize that this reflection question was referring specifically to the project guideline document rather than the annotation guidelines we developed for annotators. Although this became clear after rereading, slightly more explicit wording could help avoid that small moment of confusion.

The dataset description requirement was thorough, but there was some ambiguity regarding the expected level of technical depth. While the instructions state that the dataset should be described in enough detail to be reproducible, it is not fully specified whether a procedural explanation is sufficient or whether full reproducibility (e.g., detailed code steps) is expected. Clarifying the intended depth would help students better calibrate their responses.

Finally, although the guideline acknowledges that annotators may encounter edge cases, it does not explicitly state whether label definitions are expected to remain fixed throughout the process or whether iterative refinements are acceptable as long as they are documented. Clarifying this expectation would provide additional transparency about how flexible the annotation scheme can be.

Overall, the instructions were comprehensive and logically organized. However, adding small clarifications around terminology, procedural application, and expectations would make the assignment even clearer and further reduce uncertainty during implementation.