**# V1 Mode Output**

<!-- # Laptop Male Software Engineer Cycling Team

# <NUMBER>-<PHONE>

# Email: <EMAIL> | Twitter: @Cycling\_Team

# Cyclists are invited to participate in the Cycling Cycling World Cup in Brazil.

# The Cycling Federation of Brazil is a global organization dedicated to promoting cycling

# and the development of sustainable living. Cycling is an international organization that promotes the health,

# safety and well-being of all people. The Cycling Foundation is dedicated solely to the pursuit of the highest quality

# and quality of life -->

**# V2 Detailed Output**

Laptop Male Software Engineer Cycling Club

URL URL URL URL

\_\_\_\_\_

This e-mail message is intended only for the individual or entity to which it is addressed and may contain information that is privileged, confidential, or exempt from disclosure under applicable Federal or State law. If the reader of this e-mai-mail is not the intended recipient, or the employee or agent responsible for delivering the message to the intended recipients, you are hereby notified that any dissemination, distribution or copying of this communication is strictly prohibited. If you

**--------------------**

**--------------------**

**--------------------**

Mobile Female Data Scientist Hiking in the US

URL

--URL This message was sent by: BBC Worldwide Americas

Avenue of the Americas New York, NY,

This email is to provide you with a personal digital file of all emails from BBC Worldwide into your inbox. If you prefer not to continue receiving email communications, please unsubscribe here instead of replying to this email.

To update your profile and customize what email alerts and newsletters you receive, please click here.

Having

**--------------------**

**--------------------**

**--------------------**

Desktop Male Graphic Designer Painting -.

URL URL URL URL #

>

> -Houzz Logo

> //RESERVATION STAMPED HANDS AND LINE BLUES CURRENCY.

Image Credit: Kiki Bentonka

This is a photo message from my Houzz ideabook.

If you are having trouble viewing this, click here.

Share This: URL URL

URL

Thank You,

**--------------------**

**--------------------**

**--------------------**

**# V2 Pipeline Output**

Hello, Following up on the bubblegum shipment.

I'll send you the info.

Hope you are well.

Best,

Janeh.

Janeh.

On Tue, Feb , at : AM, Cavanaugh, Kristin < wrote:Hi Janeh,I hope you are well.I spoke to Tom Rothman and he said he'd give me a call regarding your order, I told him I wanted to talk to you about your order and the order and I said I would send you a separate email that was so that I could communicate the order.I don't know if you had a chance to look at it and he suggested I send it to him directly. I've enclosed the order and my contact info for your reference.Thank you.Best,KristinErnest,Thank you for purchasing your itemPlease note that it has arrived in your carrier.

The items listed in this shipment are:x x x inch-wide x inch-wide x inch-wide x inch-wide x inch-wide x inch-wide

**--------------------**

**--------------------**

**--------------------**

Please confirm the delivery date for our next order.

If you have any questions, please contact the Global Service Desk or your Local IT Representative.

Global Service Desk Contact Information:

North America:

US Toll Free: -SPE-SONY

Europe: ()-International Toll Fee Numbers

OnNet: - or -

GSD Live Chat

Regards,

SPE Identity Management

MP/WPF

ASAP

CRB(Competitive Releases)

DICER

FCM

GPMS - CopyRight

GPMS - MAGIC

GPMS - SCRY

GPMS - Titles and Registration

IntSales

Motion Pictures Portal

Script Tracker

SpiritWorld

Superbad

Worldwide Print Tracking System (WPTS)

Worldwide Publicity Website

Productions

C

Calypso

Dropzone

GPAS

Motion Pictures Production Database(MPPDB)

Tview

TV

BB

CC

Carmen

DealTracker

Dr. Oz

DTSM

ITSM/SARA

**--------------------**

**--------------------**

**--------------------**

Can you update me on the status of the invoice?

If not, please let me know.

Thanks,

Larry

Larry Marino

On Apr , , at : AM, Larry Marino < wrote:

Hi Larry,

Jane was hoping to see if we could get an invoice from ECS for this month's sales tax. We are working on a deal to make this work; the invoice is supposed to be CURRENCY,. I'm sorry but we won't be able to get a hard copy.

Thanks again!

Larry Marino

On Apr , , at : AM, Larry Marino < wrote:

Thanks Larry,

I hope that you are well.

I wanted to follow up with you briefly regarding the payment for the June invoice. I have received an invoice from ECS to make certain that we were billed by ECS for this year.

The invoice is supposed to be CURRENCY,.

The invoice is for May through and the invoice will be CURRENCY,.

Thanks again,

**--------------------**

**--------------------**

**--------------------**

# # Generate text using the pipeline

# email\_text1 = generate\_email("Hello, Following up on the bubblegum shipment.", use\_pipeline=True)

# email\_text2 = generate\_email("Please confirm the delivery date for our next order.", use\_pipeline=True)

# email\_text3 = generate\_email("Can you update me on the status of the invoice?", use\_pipeline=True)

# # Print the generated emails

# print(email\_text1)

# print("--------------------")

# print("--------------------")

# print("--------------------")

# print(email\_text2)

# print("--------------------")

# print("--------------------")

# print("--------------------")

# print(email\_text3)

# print("--------------------")

# print("--------------------")

# print("--------------------")

**# Result Verison 1**

Software Development Team Email

Brief project update email for a software development team meeting this weekend was provided below in advance of Tuesday's meeting. We understand that the schedule for this meeting has changed. I wanted to get everyone an update on our progress. Please provide your feedback by this Thursday, May th after we have been completed and finalized. I wish you all the best in the best possible future. Best regards, Maxene A. Verma | Marketing & Research Paralegal | COMPANY Pictures Releasing International, West Washington Blvd., Thalberg Bldg. | Culver City, CA ( | | \* Attachments: ( Bytes) x ( Bytes) x ( Bytes) x ( Bytes) ( Bytes) x ( Bytes) ( Bytes) x ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) x ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) m ( Bytes) x ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) Each individual must participate in the "Join Together" component of the Round Up Process. The Rules will be reviewed with the Rules by the Committee in partnership with the Chamber. Please contact me if you have any questions about these materials. Regards, Maxene A. Verma | Marketing & Research Paralegal | COMPANY Pictures Releasing International, West Washington Blvd., Thalberg Bldg. | Culver City, CA ( | | \* Attachments: x ( Bytes) x ( Bytes) x ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) m ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) ( Bytes) (

**--------------------**

Financial Report Request Email (with pipeline):

Detailed financial report request by a senior analyst for a Northwest bank, the company is currently conducting an online survey of its employees and business operations. Our goal is to identify, among other things, companies that have "met Wall Street's expectations" or "the market needs that are needed," the information being gathered from the survey should be deemed complete. An objective of the survey was that we would send an electronic resume, either verbal or by email in the next few days. To this end, the deadline for the survey has been extended, starting in March, to December February and ending in April th February. This will take place as follows: EES will conduct the survey using the online questionnaire to date. It will focus on identifying the most market-sceptical "attributable" employees. The results will be disclosed on Monday, with the second survey being conducted early next week, and our input will be considered during the second survey in the coming weeks. Regards, COMPANY Williams Associate Director, Public Relations If you are contacted regarding the above survey, please e-mail If you would like to be removed from future survey distribution, please e-mail If you would like to be removed from future surveys, please Please be advised that your direct reports will not be given again until completed. Thank you, Ingrid O'Donnell, Senior Manager, Commercial Support EES Business Development Contact Information You can also send an electronic resume, if you would like to have them taken a look on their qualifications for positions. Sincerely, The Survey Company DataSite Team - Customer Relationship Studies Team - POLITICAL PARTYic National Committee Montgomery Street, Suite LOCATION, Texas USA Telephone: (U) - Fax: (M) - Site Visit Map:<(toll-free):<(filing location: Hyatt Regency, Orange County, CA) Thanks for your participation and participation on the survey! Sincerely, The Survey Company DataSite Team - Customer Relationship Studies Team - POLITICAL PARTYic National Committee Montgomery Street, Suite LOCATION, Texas USA Telephone: (U) - Fax: (M) - Site Visit Map:<(toll-free):<(filing location: Hyatt Regency, Orange County, CA) All e-mails come directly from The Survey Company. In addition to being members of the Company's Executive Team who are not on the Company's Executive Team please consider the environment before printing this e-mail. This message was sent to: If you no longer wish to receive these emails, click on the following link: Unsubscribe. <a style="text-align:center!important; width="px!important; }.ExternalClass p:link, span.intent, div.us\_SignInRow {margin: px ;} .ExternalClass font { font-family:"Calibri", "Verdana", "Helvetica", line-height:px; padding: px!important;} @media only screen and (max-width: px) { \*.layout-, \*.footer-layout { width:px!important; } \*.layout-, \*.center-align, \*.show-icons,\*.one-column, \*.mediumgrid-header, \*.outer-layout, \*.media-layout, \*.grid-layout, \*.header-layout, \*.featureagentads-layout, \*.ads-layout, \*.housead-layout, \*.app, \*.subscribe, \*.unlimited-access, \*.footer, \*.access, \*.app-class, \*.app-span, \*.copyright-block, \*.view-layout { width:px!important; } \*.media-layout { padding: px!important; margin:!important; } \*.view-layout { margin: px !important; } \*.grid-layout { margin:px px px!important; } \*.header-layout

--------------------

Marketing Campaign Creative Brief (with pipeline):

Creative brief for a new marketing campaign targeting young adults and their fans: It's not clear what the plan is for this animated movie from the beginning. ) How to Give an EXTRA EXPENSION of the trailer with the promise of a new brand, it's going to work very well for us - especially as in the new trailer for "Grown Ups ," but we need to have a feeling about how much fun and powerful "Grown Ups " has been. We will be discussing this with everyone this weekend. ) The new trailer for Interview # (which is still in production) promises a lot more emotion and action for the film. Cc: Vollack, Lia; Marshall, Bob; Guerin, Jean This wasn't made in-the-works by Mike. They'd also like to give the current joke - an old joke, without any real conflict, to the movie, without the promise that it's a sequel or something. The script (and script) incorporates language I understand from the conversations with James and Phil last week: Regarding Interview # we just wanted you to know that we are committed to working on all the points on the horizontal and vertical artwork to show off COMPANY's upcoming campaign trailer. We want to address a number of concerns namely, if the'story' is used incorrectly, it's no good to use references from the previous trailers. If this isn't clear, then we may want to adjust and clarify what we mean by using some additional language to the story/concept. As we talked on Friday, I'm sending you the following email (from JLL) regarding several instances where we were able to bypass out our own campaign-include one that's more generic. But again, we will need to communicate with JLL about these issues/ concerns so that they don't pop into places where we are literally communicating. If this isn't clear then we may want to go ahead with the other campaign to get us next steps. Hereto be clear: when we talk about the possibility of splitting up the various businesses (Digital, Media Networks, Home Entertainment etc.), we understand the scope of the partnership, but are concerned about different businesses that can serve multiple brands in one fashion or another. We don't want to appear too stupid or wrong, however when we talk about a partnership that could deliver big gains in the long run. Again, this is confidential, and we will need to ask JLL to help provide the answer asap. On May , , at : PM, "Belgrad, Doug" < wrote: Hey guys, Just got off the phone with JLL about the below-referenced positioning for the Playstation in the wake of today's announcement. We're excited to be collaborating with the Alliance and Disney on Uncharted and Bloodshot, both coming out before The Avengers. JLL and the Alliance have made a tremendous impact in terms of funding for the three games (the "Smurfs" series, Venom and Sinister Six). There are some exciting new features that these are potential tentpoles in the future: Battle of the Five - A look at our two game plans by Shane Black's team as well as looking at how Activision can benefit from this in developing four separate game based games and whether he is leveraging the technology here. Booming In Productions - We want to do an ambitious and successful pitch to support the vision of our motion picture assets. We feel that this approach will deliver big returns for us. The biggest difference from our pre-COMPANY Pictures strategy was with the release of Rise of Electro, which cost more than CURRENCY million at D (a jump from Budget). We are currently working on many more franchises, including Bond franchise, Dr. Seuss, Batman and more. We're also expanding the Spider-Man universe and expanding the world of DC with a darker tone and feel for Spider-Man. Additionally, we would like to get a sense of what COMPANY can and can't hide and what we're working on to achieve in those three games. Hoping you all have a good weekend. Doug On May , , at : PM, "Gumpert, Andrew" < wrote: Believe me, I am looking forward to speaking with you soon. Look forward to talking. mike....... Andrew Gumpert President, Worldwide Business Affairs & Operations COMPANY Telephone: Facsimile: E-Mail: PRIVACY NOTICE: ument by mistake, please e-mail the sender securely dispose of it.

**--------------------**

Software Development Team Email (without pipeline):

Brief project update email for a software development team, please contact me. I look forward to hearing from you. If you have any questions, you can reach me at . I will be in meetings this afternoon. Thanks so much! Best regards, Shelly SHALINIE BERMAN | CONTACT | SOLUTIONS INC | P: | F: () Email: Please see attached draft correspondence from the SEC on file sharing. We have also discussed the possibility of sending around an email that shows marked "Full Schedule". We are scheduled to meet as soon as possible, but would love to have you guys review. My team has also gotten some comments from Kaz Hirai and want to make sure this is covered as well. Thank you all very much for your cooperation. Best, - Sharon > > > *\_\_*\_\_\_ This e-mail and any files transmitted with it are intended solely for the use of the individual or entity to whom they are addressed. This e mai has been sent to the attention of only the named addressee(s). If the reader of this e mail is not the intended recipient or the employee or agent responsible for delivering the message to thi you are hereby notified that any use dissemination, forwarding, printing or copying of th e com e of any sort is strictly prohibited. Any views or opinions expressed are solely those of sakethe only author and do not necessarily represent those or should necessarily be taken as an excuse. e \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ SUBSCRIBE TO OUR AWARD WINNING NEWSLETTERS TRAVEL TIPS, THE WINE LIST AND THE DAILY FOOD & Wine THE DISH NUM

**--------------------**

Financial Report Request Email (without pipeline):

Detailed financial report request by a senior analyst that includes related companies. The analyst estimates that the analyst's annual risk forecast will be CURRENCY billion by July S. The analyst believes this year will reach C- million. As such, it was profitable to report that number from NUMABS and other non-financial companies while also operating as margin and basis. He estimated that COMPANY may see a net loss of CMM, which would hit COMPAE, or C/b. If all alternatives are not feasible, then he expects his team to continue through NUMP and report to each of them this November. After that, management thinks the stock should close to its October start date. And COMPENSENSON is preparing additional capital expenditure in the coming months, according to Tomn Bersch. Berszch was also informed last year that at the end of FYNUMBIRING THE SECURITY RANGE, NUMRBIO would require more capital, because of this, had COMPensenRON's business practices reported on their website last September. That was more accurate than the sum sheet in October that came out next week. His team made clear they did not want a "crashed" release as a result. At the same time, Berszlch also admitted that some people have assumed the public can view COMPenron's assets in a certain light. Below is his thoughts. According to an internal memo sent to me by the SEC this afternoon, no one from COMP's group has responded via email to us with the numbers that were not included in our analysis. I think what I will assume, however, is that we will put on solid media tomorrow. Hopefully that nothing leaks before Friday. To my credit, you probably already know that there are several possible investors this fiscal year - including several senior COMPanyer named as potential investment bankers. All of the above is considered a necessary investment. However, in order to gain a timely and accurate portrayal of our overall portfolio, we are going to rely heavily on media coverage throughout the process. When we do look at what media is covering, I believe that this proceeds will likely flow as we work closely with both sides to understand what the opportunities are and what types of questions we're looking for;

**--------------------**

Marketing Campaign Creative Brief (without pipeline):

Creative brief for a new marketing campaign targeting young adults + -Minute TV Spot for First Time ABC-Disney's 'The Walking Dead' (Video) o It looks like ABC has the right to promote the show on a big screen. Here are a few highlights: The Young Adult Swim, Disney Channel and ABC Family were among the top rated networks in the last year of its second season Digital TV Spots for st Century Fox, NBCU Cable, FOX Sports Network and Yahoo! Television announced last week that they will be joining the network along with their networks' digital ad network, Nickelodeon and VHNUMB. "The idea of creating video content in order to drive revenue through the process," said Doug Belgrad, president of worldwide commercial for COMPANY. "This is the first time in our history that a network could have an effect on ad viewing. " In addition, AOL launched video streaming service VuduStream in February. The service allows subscribers to watch live sports videos on their mobile devices. Vubu is a major player in that industry, bringing in more than NUMB of pay television subscribers from around the world to view and comment on content across their homes. Ad-supported services also allow viewers to stream live sporting events across multiple platforms. About VUDuSource presents a powerful advertising strategy. For more information please click here. Contact Lisa Woo at to schedule a phone interview or email. This e-mail and any files transmitted with it are intended solely for the use of the individual or entity to whom they are addressed. If the reader of this e E-Mail is not the intended recipient or the employee or agent responsible for delivering the message to the original recipient, you are hereby notified that any use dissemination, forwarding, printing or copying of thi Attachments: [imageNUMb. A Letter from the Publisher and CEO of Vodu Digital Distribution, Inc. vid. This electronic file (including any attachments) contains information concerning VOD and/or electronic sell-through (EST), content protection and other technology related to VAN. IF YOU HAVE RECEIVED THIS COMMUNICATION IN ERROR, PLEASE NOTIFY US IMMEDIATELY BY TELEPHONING THE ORIGINAL COMMUNITY AND DESTROY ALL COPIES, BOTH ELECTRONIC AND OTHER, OF

**--------------------**

**# Result Verison 2**

Software Development Team Email

Brief project update email for a software development team meeting this weekend was provided below in advance of Tuesday's meeting. We understand that the schedule for this meeting has changed. I wanted to get everyone an update on our progress. Please provide your feedback by this Thursday, May th after we have been completed and finalized. I wish you all the best in the best possible future. Best regards, Maxene A. Verma | Marketing & Research Paralegal | COMPANY Pictures Releasing International, West Washington Blvd., Thalberg Bldg. | Culver City, CA ( | |

**--------------------**

**# Result Verison 3**

Software Development Team Email

Brief project update email for a software development team meeting this weekend was provided below in advance of Tuesday's meeting. We understand that the schedule for this meeting has changed. I wanted to get everyone an update on our progress. Please provide your feedback by this Thursday, May th after we have been completed and finalized. I wish you all the best in the best possible future. Best regards, Maxene A. Verma

**--------------------**

**# 使用pipeline生成文本**

custom\_text1\_with\_pipeline = generate\_email("Brief project update email for a software development team", use\_pipeline=True)

custom\_text2\_with\_pipeline = generate\_email("Detailed financial report request by a senior analyst", use\_pipeline=True)

custom\_text3\_with\_pipeline = generate\_email("Creative brief for a new marketing campaign targeting young adults", use\_pipeline=True)

**# 不使用pipeline生成文本**

custom\_text1\_without\_pipeline = generate\_email("Brief project update email for a software development team", use\_pipeline=False)

custom\_text2\_without\_pipeline = generate\_email("Detailed financial report request by a senior analyst", use\_pipeline=False)

custom\_text3\_without\_pipeline = generate\_email("Creative brief for a new marketing campaign targeting young adults", use\_pipeline=False)

**# 打印使用pipeline生成的文本**

print("Software Development Team Email (with pipeline):")

print(custom\_text1\_with\_pipeline)

print("--------------------")

print("Financial Report Request Email (with pipeline):")

print(custom\_text2\_with\_pipeline)

print("--------------------")

print("Marketing Campaign Creative Brief (with pipeline):")

print(custom\_text3\_with\_pipeline)

print("--------------------")

**# 打印不使用pipeline生成的文本**

print("Software Development Team Email (without pipeline):")

print(custom\_text1\_without\_pipeline)

print("--------------------")

print("Financial Report Request Email (without pipeline):")

print(custom\_text2\_without\_pipeline)

print("--------------------")

print("Marketing Campaign Creative Brief (without pipeline):")

print(custom\_text3\_without\_pipeline)

print("--------------------")

# Basic cleanup: remove redundant phrases and unwanted tokens

# text = re.sub(r'\b(URL)\b\s\*', '', text) # Remove 'URL' placeholders

# text = re.sub(r'\b(\d{1,2}[:]\d{2}\s\*(AM|PM))\b', '', text) # Remove standalone time

# text = re.sub(r'\[\d+\]', '', text) # Remove citation-like numbers

# text = re.sub(r"NUMBER|EMAIL|PHONE|FAX", "", text)

# text = re.sub(r"\[.\*?\]|-\s\*-\s\*|<.\*?>", "", text) # Clean up brackets and dashed lines

# text = re.sub(r'\s{2,}', ' ', text) # Replace multiple spaces with a single space

# text = re.sub(r'\\* Attachments:.\*', '', text, flags=re.S) # 删除'Attachments:'及其后的所有内容

# text = re.sub(r'\bAttachments:\s\*(\( Bytes\)\s\*)+(x\s\*\( Bytes\)\s\*)\*', '', text, flags=re.S)

# text = re.sub(r'\|\s.\*', '', text) # Remove suffixes after names # This regex removes everything after a '|' character

from transformers import pipeline, AutoModelForCausalLM, AutoTokenizer, set\_seed

import re

import spacy

**# Load Spacy's English tokenizer**

nlp = spacy.load('en\_core\_web\_sm')

**# Set a random seed for reproducible results**

set\_seed(42)

**# Load the model and tokenizer**

model\_checkpoint = "./saved\_model"

tokenizer = AutoTokenizer.from\_pretrained(model\_checkpoint, use\_fast=True, padding\_side='left')

if tokenizer.pad\_token is None:

tokenizer.add\_special\_tokens({'pad\_token': '[PAD]'})

model = AutoModelForCausalLM.from\_pretrained(model\_checkpoint)

**# Set up the pipeline using the freshly trained model and tokenizer**

generator = pipeline('text-generation', model=model, tokenizer=tokenizer)

def advanced\_text\_cleaning(text):

# Use Spacy to parse sentences and filter out overly short fragments and placeholders

doc = nlp(text)

cleaned\_text = []

for sent in doc.sents:

# Filtering out too short sentences and placeholders

if len(sent.text) > 20 and not re.search(r'NUMB[RL]|NUMBE[RSZ]', sent.text):

cleaned\_text.append(sent.text.replace('\n', ' ').strip())

return ' '.join(cleaned\_text)

def clean\_generated\_text(text):

return advanced\_text\_cleaning(text.strip())

def generate\_email(product, gender, profession, hobby, use\_pipeline=False):

# 构建输入文本

input\_text = f"As a {gender.lower()} {profession} interested in {hobby}, I am looking for a {product} that suits my needs."

if use\_pipeline:

result = generator(input\_text, max\_length=1024, do\_sample=True, top\_k=50, temperature=0.9, repetition\_penalty=1.1, truncation=True)

generated\_text = result[0]['generated\_text']

else:

inputs = tokenizer(input\_text, return\_tensors="pt", max\_length=512, truncation=True, padding="max\_length")

outputs = model.generate(

\*\*inputs,

max\_length=1024, # Increase maximum length

temperature=0.9, # Higher diversity

top\_k=50, # Broader vocabulary choice

top\_p=0.95, # Nucleus sampling

no\_repeat\_ngram\_size=2, # Allow minimal repetition

repetition\_penalty=1.1 # Slightly less repetition penalty

)

generated\_text = tokenizer.decode(outputs[0], skip\_special\_tokens=True)

return clean\_generated\_text(generated\_text)

**# 示例参数**

product = "Laptop"

gender = "Male"

profession = "Software Engineer"

hobby = "Gaming"

**# 生成电子邮件**

email\_text = generate\_email(product, gender, profession, hobby, use\_pipeline=True)

print("Generated Email:")

print(email\_text)

print("--------------------")

Generated Email:

Laptop Male Software Engineer Gaming Technology - COMPANY Corporation of America (NYSE: SNE)

Experience the power and excitement that drives gaming, social games & entertainment. Experience your personalized experiences with a variety of mobile devices from anywhere in our network including tablets to smartphones, game consoles, plasma TVs, Blu-ray players etc.. The experience is immersive so you can experience all aspects at one time while exploring other technologies. You will also be able to access exclusive content such as Pandora's World which allows users to listen directly on their phones or watch videos instantly on any device. We are confident this will help us achieve more than NUMBER million people worldwide who have already purchased product via apps like Facebook Messenger, Twitter API for Android and Windows Phone/iPad Air Connectivity. For more information please visit URL

URL logo

Questions? Call Igor Gruppin PHONE Fax +PHONE EMAIL

To contact Member Services | View Our Privacy Statement | Unsubscribe

©️NUMBER Microsoft Corporation. All rights reserved. | Acceptable Use Policy | Contact Customer Service | NUMBER-HOUR CONTROL CENTER| USA

This email was sent by: Microsoft Corporation

NUMBER Marta St., Suite NUMBERSan Diego del Norte, CA NUM

Generated Email:

Laptop Male Software Engineer Gaming Technology - COMPANY

Experience the power of mobile devices and make sure you are comfortable with your smartphone.

Mobile device Management is a key component to achieving success in our industry, helping us achieve breakthroughs for business leaders by leveraging the best-in-class experience available. We have been working closely together on numerous projects that will benefit from this evolution into an integrated marketing tool used throughout all aspects including media creation & distribution; digital content development/retail fulfillment (FOD), social networking forums such as Twitter, Facebook etc.; strategic communications solutions like Magento CMS or even SAPPRFT products.

We would be honored if you could join us at one another's pace! Please let me know what time works best for everyone else who needs it most especially during their busy schedules. Thank you so much for considering this request. I look forward to hearing from you soon. Best regards, Karen Kondo | Senior Vice President Corporate Social Responsibility | COMPANY Corporation of America NUMBER Madison Avenue, Room NUMBERSNUMBERCulver City, CA NUMBER( PHONE| EMAIL)

Attachments:

x (PHONE Bytes)

ATTNUMBER.htm ("");<EMAIL

X (PHONE By

**--------------------**

# Basic cleaning

# text = re.sub(r'^(Re:|Fwd:)', '', text) # Remove reply and forward marks

# text = re.sub(r'Best regards,.\*$', '', text, flags=re.DOTALL) # Remove everything after signature

# text = re.sub(r'PHONE.\*$', '', text, flags=re.DOTALL) # Remove everything after phone numbers

# text = re.sub(r'Email:.\*$', '', text, flags=re.DOTALL) # Remove everything after email addresses

# text = re.sub(r'Cc:.\*$', '', text, flags=re.DOTALL) # Remove CC list

# text = re.sub(r'\\* Attachments:.\*', '', text, flags=re.S) # Remove 'Attachments:' and everything following it

# text = re.sub(r'©️ .\*$', '', text, flags=re.DOTALL) # Remove copyright and ownership statements

# text = re.sub(r'URL If this message is not displaying properly, click here.\*$', '', text, flags=re.DOTALL) # Remove error display message and links

# text = re.sub(r'\d{5,}', 'NUMBER', text) # Replace long sequences of numbers, likely phone numbers or ZIP codes

def generate\_email(product, gender, profession, hobby):

input\_text = f"{product} {gender} {profession} {hobby}"

result = generator(

input\_text, # Initial text to prompt the model. Sets the context or topic for text generation.

max\_length=512, # Maximum length of the generated text in tokens, limiting the output size.

do\_sample=True, # Enables stochastic sampling; the model can generate diverse outputs at each step.

top\_k=20, # Limits the vocabulary considered at each step to the top-k most likely next words.

top\_p=0.6, # Uses nucleus sampling: Narrows down to the smallest set of words totaling 60% of the likelihood.

temperature=0.4, # Scales logits before sampling to reduce randomness and produce more deterministic output.

repetition\_penalty=1.5, # Penalizes words that were already mentioned, reducing repetition in the text.

truncation=True, # Truncates the output to the maximum length if it exceeds it.

num\_return\_sequences=3 # Generates three different sequences to choose from, enhancing output variety.

)

# Select the best output from the generated sequences

best\_text = sorted([clean\_generated\_text(r['generated\_text']) for r in result], key=len)[-1]

return best\_text

def generate\_email(product, gender, profession, hobby):

input\_text = f"{product} {gender} {profession} {hobby}"

result = generator(

input\_text, # The starting text that guides the model on what to generate.

max\_length=256, # Limits the length of the generated text to 512 tokens.

truncation=True, # Cuts off the text if it goes beyond the maximum length.

top\_k=20, # Considers only the top 20 most likely next words at each step.

top\_p=0.6, # Limits the word choices to the top 60% most likely options.

temperature=0.4, # Lowers randomness to make the output more predictable.

repetition\_penalty=1.5, # Prevents the model from reusing previously mentioned words.

num\_return\_sequences=2, # Produces two versions of generated text.

do\_sample=True, # Enables stochastic sampling; the model can generate diverse outputs at each step.

)

def clean\_generated\_text(text):

#Basic cleaning

text = re.sub(r'^(Re:|Fwd:)', '', text) # Remove reply and forward marks

text = re.sub(r'Best regards,.\*$', '', text, flags=re.DOTALL) # Remove everything after signature

text = re.sub(r'PHONE.\*$', '', text, flags=re.DOTALL) # Remove everything after phone numbers

text = re.sub(r'Email:.\*$', '', text, flags=re.DOTALL) # Remove everything after email addresses

text = re.sub(r'Cc:.\*$', '', text, flags=re.DOTALL) # Remove CC list

text = re.sub(r'\\* Attachments:.\*', '', text, flags=re.S) # Remove 'Attachments:' and everything following it

text = re.sub(r'©️ .\*$', '', text, flags=re.DOTALL) # Remove copyright and ownership statements

return text.strip()

# text = re.sub(r'Best regards,.\*$', '', text, flags=re.DOTALL) # 移除签名后的内容

# text = re.sub(r'©️ .\*$', '', text, flags=re.DOTALL) # 移除版权和所有权声明