

"""	pan_numbers =	
Assignment No 5	re.findall(pan_number_pattern, text)	# OUTPUT -
Name – Ishwar wagh		"""
Batch – B4	# Extract spaCy entities	URLs: ['https://www.example.com.']
Roll No - 65	entities = [(ent.text, ent.label_) for ent in doc.ents]	IP Addresses: ['192.168.1.1']
Assignment Title: Implement regular expression function to find URL, IP address, Date, PAN number in textual data using python libraries	return {	Dates: ['2023-01-01']
"""	'urls': urls,	PAN Numbers: ['ABCDE1234F']
	'ip_addresses': ip_addresses,	Entities: [('IP', 'ORG'), ('192.168.1.1', 'CARDINAL'), ('2023-01-01', 'DATE'), ('PAN', 'ORG')]
import spacy	'dates': dates,	"""
import re	'pan_numbers': pan_numbers,	
	'spaCy_entities': entities	
	}	
# Load the spaCy English language model		
nlp = spacy.load("en_core_web_sm")	# Example usage	
	text_data = """	
# Define regular expressions	Here is a sample text with a URL: https://www.example.com.	
url_pattern = re.compile(r'https?://\S+ www\.\S+')	Also, an IP address: 192.168.1.1. The date is 2023-01-01,	
ip_address_pattern = re.compile(r'\b(?:\d{1,3}\.){3}\d{1,3}\b')	and a PAN number is ABCDE1234F.	
	"""	
date_pattern = re.compile(r'\d{4}- \d{2}-\d{2}')	results = extract_entities(text_data)	
pan_number_pattern = re.compile(r'[A-Z]{5}[0-9]{4}[A-Z]')	print("URLs:", results['urls'])	
	print("IP Addresses:", results['ip_addresses'])	
def extract_entities(text):	print("Dates:", results['dates'])	
# Tokenize the text using spaCy	print("PAN Numbers:", results['pan_numbers'])	
doc = nlp(text)		
# Find entities using regular expressions	print("Entities:", results['spaCy_entities'])	
urls = re.findall(url_pattern, text)		
ip_addresses = re.findall(ip_address_pattern, text)		
dates = re.findall(date_pattern, text)		