Sentiment Analysis with Azure Cognitive Service

By:

Agus Suparno, S.Si., M.Eng., MVP.



Agus Suparno, S.Si., M.Eng., MVP





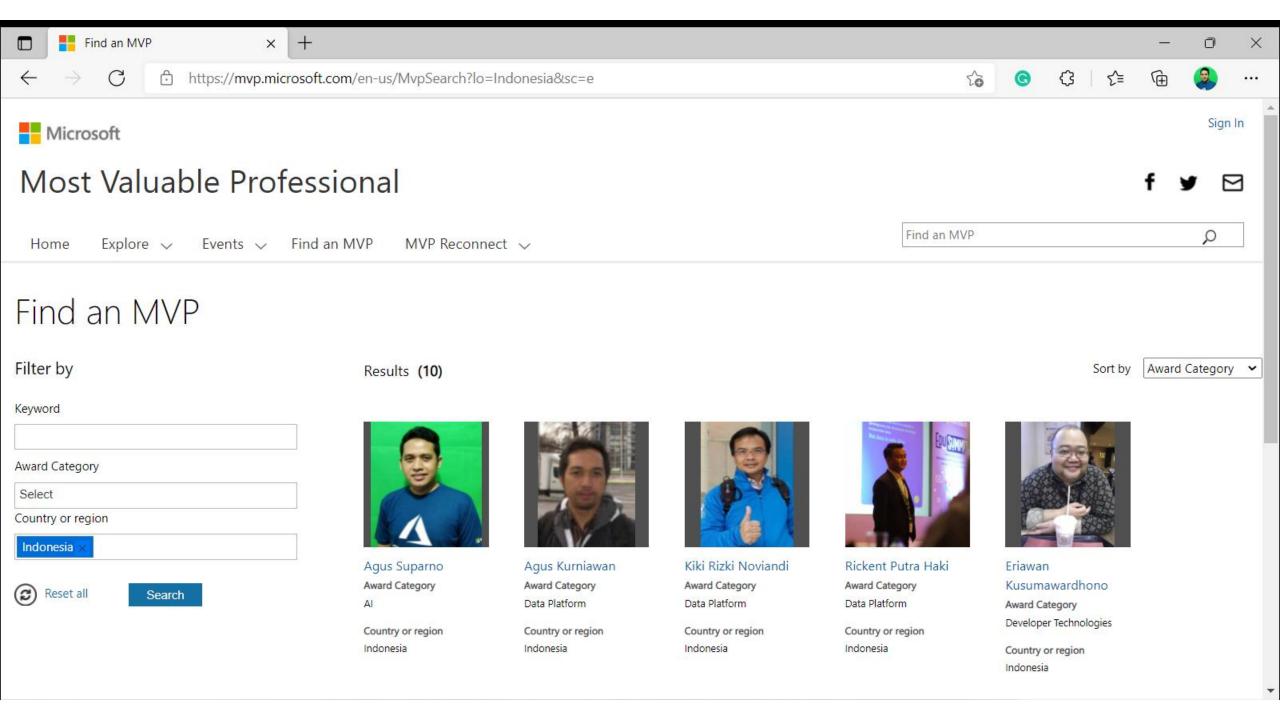


Microsoft MVP, Leader of Mugi Purwokerto, Lecturer, CEO of Sakti Builder









Artificial Intelligence

Artificial Intelligence



Any technique that enables computers to mimic human intelligence. It includes machine learning

Machine Learning



A subset of Al that includes techniques that enable machines to improve at tasks with experience. It includes deep learning

Deep Learning



A subset of machine learning based on neural networks that permit a machine to train itself to perform a task.

Azure Al

- Azure Applied Al Service
- Azure Cognitive Service
- Azure Machine Learning
- Al infrastructure

#1 Azure Applied Al Service

Specialized AI services for specific business scenarios

- Modernize business processes with taskspecific AI to solve for common scenarios.
- Accelerate development with built-in business logic that enables you to launch solutions in days—not months.
- Run responsibly anywhere with security that extends from the cloud to intelligent edge.



#1 Azure Applied Al Service

CONVERSATIONS



Azure Bot Service

Create bots and connect them across channels.

DOCUMENTS



Azure Form Recognizer

Turn documents into usable data at a fraction of the time and cost.

SEARCH



Azure Cognitive Search

Bring Al-powered cloud search to your mobile and web apps.

MONITORING



Azure Metrics Advisor

Proactively monitor metrics and diagnose issues.

VIDEOS



Azure Video Analyzer

Extract actionable insights from your videos.

ACCESSIBILITY



Azure Immersive Reader

Help users read and comprehend text.

A comprehensive family of customizable cognitive APIs for vision, speech, language, and decision making

- Easily access sophisticated AI models, including OpenAI models, with the most comprehensive portfolio of AI capabilities on the market.
- **Build confidently** with the first AI services to achieve human parity in computer vision, speech, and language.
- **Deploy anywhere** from the cloud to the edge with containers.
- **Get started quickly** and customize solutions to meet your own unique requirements.



Speach



- Language
- Vision
- Decision

Speech to text

Transcribe audible speech into readable, searchable text.

Text to speech

Convert text to lifelike speech for more natural interfaces.

Speech translation

Integrate real-time speech translation into your apps.

Speaker recognition

Identify and verify the people speaking based on audio

- Speach
- Language



- Vision
- Decision

Entity recognition

Identify commonly-used and domain-specific terms.

Sentiment analysis

Automatically detect sentiments and opinions from text.

QnA Maker

Distill information into easy-to-navigate questions and answers.

Conversational language understanding

Enable your apps to interact with users through natural language.

Translator

Translate more than 100 languages and dialects.

- Speach
- Language
- Vision
- Decision



Analyze content in images and video.

Custom Vision

Customize image recognition to fit your business needs.

Face API

Detect and identify people and emotions in images

- Speach
- Language
- Vision
- Decision



Identify potential problems early on.

Content Moderator

Detect potentially offensive or unwanted content.

Personalizer

Create rich, personalized experiences for every user.

#3 Azure Machine Learning

An end-to-end platform for building, training, and deploying machine learning models

- Develop with your choice of tools with Jupyter Notebook, drag-and-drop designer, and automated machine learning.
- Create and deploy models at scale using automated and reproducible machine learning workflows.
- Innovate responsibly with a rich set of built-in responsible capabilities to help you understand, protect, and control data, models and processes.
- Build your way with best-in-class support for opensource frameworks and languages including MLflow, Kubeflow, ONNX, PyTorch, TensorFlow, Python, and R



#4 Al Infrastructure

Advanced, large-scale AI infrastructures and innovative training tools

- Access large-scale infrastructure with hyperclusters of thousands of state-of-the-art GPUs, providing AI accelerations that are interconnected with the latest high-bandwidth networks inside of every server.
- Enable hybrid and multi-cloud deployments with Azure Arc-enabled Kubernetes.
- Access a breadth of Al hardware with specialized hardware like FPGAs, the most comprehensive set of GPUs, and an array of general-purpose CPUs



Azure Cognitive Service – Sentiment Analysis

- Menganalisis teks adalah proses di mana Anda mengevaluasi berbagai aspek dokumen atau frasa, untuk mendapatkan wawasan tentang konten teks itu.
- Sebagai contoh, Kita mungkin membaca beberapa teks dan mengidentifikasi beberapa frase kunci yang menunjukkan poin pembicaraan utama dari teks.
- Mungkin juga kita mengenali nama orang atau tempat terkenal seperti Menara Eiffel.
- Meskipun terkadang sulit, Anda mungkin juga dapat memahami perasaan orang tersebut ketika mereka menulis teks, yang juga dikenal sebagai **sentimen**.

Text Analytics Techniques

- Text Analytics adalah proses di mana algoritme kecerdasan buatan (AI), yang berjalan di komputer, mengevaluasi atribut yang sama di dalam teks, untuk menentukan insight tertentu.
- Seseorang biasanya akan mengandalkan pengalaman dan pengetahuan mereka sendiri untuk mendapatkan insight/wawasan.
- Sebuah komputer harus dilengkapi dengan pengetahuan yang sama untuk dapat melakukan tugas tersebut. Ada beberapa teknik yang umum digunakan yang dapat digunakan untuk membangun perangkat lunak untuk menganalisis teks, antara lain:

Techniques that used to analyze text

Statistical analysis of terms used in the text.

Misalnya, menghapus "kata henti" umum (kata-kata seperti "yang" atau "sebuah", yang mengungkapkan sedikit informasi semantik tentang teks), dan melakukan analisis frekuensi dari kata-kata yang tersisa (menghitung seberapa sering setiap kata muncul) dapat memberikan petunjuk tentang subjek utama teks.

- Memperluas analisis frekuensi ke frasa multi-istilah,
 - umumnya dikenal sebagai N-gram (frasa dua kata adalah bi-gram, frasa tiga kata adalah tri-gram, dan seterusnya).
- Menerapkan algoritma stemming atau lemmatization untuk menormalkan kata sebelum menghitungnya

misalnya, sehingga kata-kata seperti "power", "powered", dan "powerful" ditafsirkan sebagai kata yang sama.

Techniques that used to analyze text

- Menerapkan aturan struktur linguistik untuk menganalisis kalimat misalnya, memecah kalimat menjadi struktur seperti percabangan seperti frasa kata benda, yang dengan sendirinya mengandung kata benda, kata kerja, kata sifat, dan sebagainya.
- Mengkodekan kata atau istilah sebagai fitur numerik yang dapat digunakan untuk melatih model pembelajaran mesin. Misalnya, untuk mengklasifikasikan dokumen teks berdasarkan istilah yang dikandungnya. Teknik ini sering digunakan untuk melakukan analisis sentimen, di mana suatu dokumen diklasifikasikan sebagai positif atau negatif.
- Membuat model vektor yang menangkap hubungan semantik antara kata-kata dengan menetapkannya ke lokasi dalam ruang n-dimensi. Teknik pemodelan ini mungkin, misalnya, memberi nilai pada kata "bunga" dan "tanaman" yang menempatkan mereka berdekatan satu sama lain, sementara kata "skateboard" mungkin diberi nilai yang memposisikannya lebih jauh.

- Meskipun teknik ini dapat digunakan dengan efek yang luar biasa, memprogramnya bisa menjadi rumit. Di Microsoft Azure, layanan "Language Cognitive Service" dapat membantu menyederhanakan pengembangan aplikasi dengan menggunakan model terlatih yang dapat:
- Tentukan bahasa dokumen atau teks (misalnya, Prancis atau Inggris).
- Melakukan analisis sentimen pada teks untuk menentukan sentimen positif atau negatif.
- Ekstrak frase kunci dari teks yang mungkin menunjukkan poin pembicaraan utamanya.
- Mengidentifikasi dan mengkategorikan entitas dalam teks. Entitas dapat berupa orang, tempat, organisasi, atau bahkan barang sehari-hari seperti tanggal, waktu, jumlah, dan sebagainya.

Contoh penerapan aplikasi

- Penganalisis umpan media sosial untuk mendeteksi sentimen seputar kampanye politik atau produk di pasar.
- Sebuah aplikasi pencarian dokumen yang mengekstrak frase kunci untuk membantu meringkas materi pokok utama dokumen dalam katalog.
- Alat untuk mengekstrak informasi merek atau nama perusahaan dari dokumen atau teks lain untuk tujuan identifikasi.

Sentiment analysis

- Kemampuan analitik teks dalam "Language Service" dapat mengevaluasi teks dan mengembalikan skor sentimen dan label untuk setiap kalimat. Kemampuan ini berguna untuk mendeteksi sentimen positif dan negatif di media sosial, ulasan pelanggan, forum diskusi, dan lainnya.
- Menggunakan model klasifikasi pembelajaran mesin yang dibuat sebelumnya, layanan mengevaluasi teks dan mengembalikan skor sentimen dalam kisaran 0 hingga 1, dengan nilai yang mendekati 1 sebagai sentimen positif. Skor yang mendekati tengah kisaran (0,5) dianggap netral atau tidak tentu.

Misalnya, dua ulasan restoran berikut dapat dianalisis untuk sentimen:

#1 "We had dinner at this restaurant last night and the first thing I noticed was how courteous the staff was. We were greeted in a friendly manner and taken to our table right away. The table was clean, the chairs were comfortable, and the food was amazing."

"Kami makan malam di restoran ini tadi malam dan hal pertama yang saya perhatikan adalah betapa sopan stafnya. Kami disambut dengan ramah dan langsung dibawa ke meja kami. Mejanya bersih, kursinya nyaman, dan makanannya enak. sangat menakjubkan."

Misalnya, dua ulasan restoran berikut dapat dianalisis untuk sentimen:

#2. "Our dining experience at this restaurant was one of the worst I've ever had. The service was slow, and the food was awful. I'll never eat at this establishment again."

"Pengalaman bersantap kami di restoran ini adalah salah satu yang terburuk yang pernah saya alami. Layanannya lambat, dan makanannya enak. Saya tidak akan pernah makan di restoran ini lagi."

Skor sentimen untuk ulasan pertama mungkin sekitar 0,9, menunjukkan sentimen positif; sedangkan skor untuk review kedua mungkin mendekati 0,1, menunjukkan sentimen negatif.

Indeterminate sentiment

- Skor 0,5 mungkin menunjukkan bahwa sentimen teks tidak dapat ditentukan, dan dapat dihasilkan dari teks yang tidak memiliki konteks yang cukup untuk membedakan sentimen atau frasa yang tidak memadai. Misalnya, daftar kata dalam kalimat yang tidak memiliki struktur, dapat menghasilkan skor yang tidak pasti.
- Contoh lain di mana skor mungkin 0,5 adalah dalam kasus di mana kode bahasa yang salah digunakan. Kode bahasa (seperti "en" untuk bahasa Inggris, atau "fr" untuk bahasa Prancis) digunakan untuk memberi tahu layanan tentang bahasa yang digunakan teks tersebut. Jika Anda mengirimkan teks dalam bahasa Prancis tetapi memberi tahu layanan tersebut kode bahasanya adalah en untuk bahasa Inggris, layanan akan mengembalikan skor tepat 0,5

Key phrase extraction

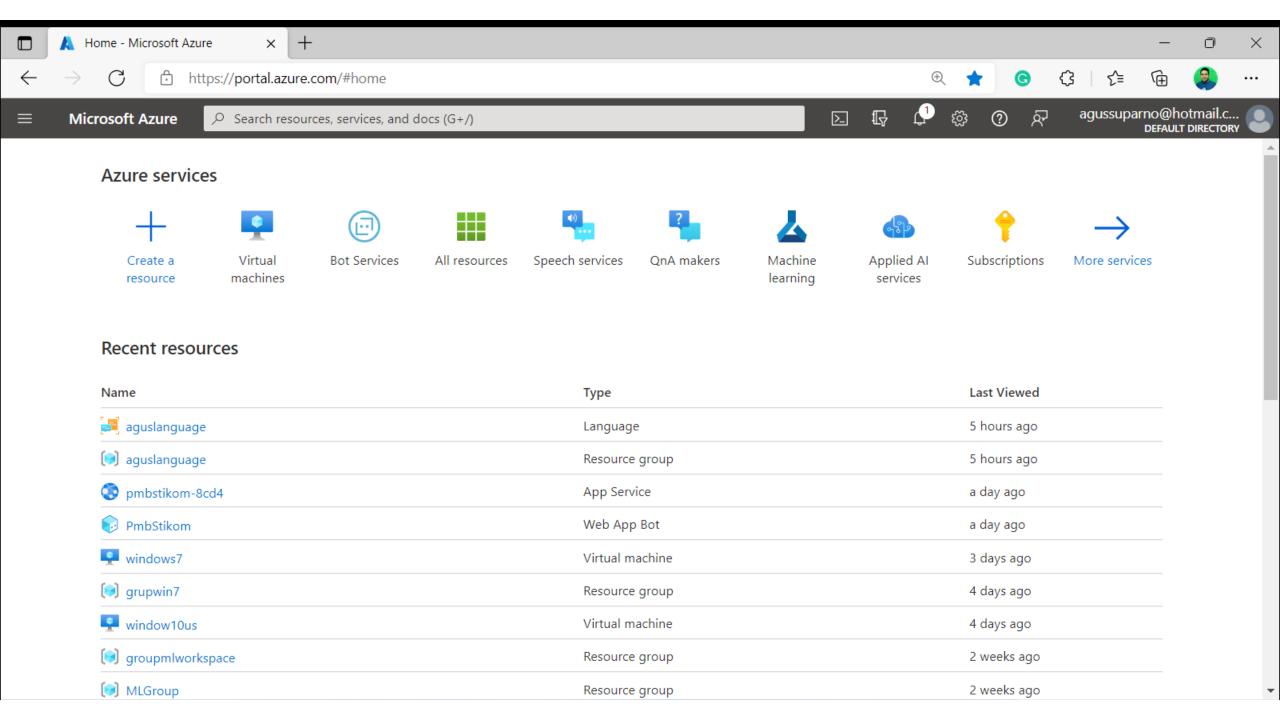
- Ekstraksi frase kunci adalah konsep mengevaluasi teks dokumen, atau dokumen, dan kemudian mengidentifikasi poin pembicaraan utama dari dokumen (s).
- Berdasarkan skenario restoran yang dibahas diatas. Bergantung pada volume survei yang telah Anda kumpulkan, mungkin perlu waktu lama untuk membaca ulasan. Sebagai gantinya, Anda dapat menggunakan kemampuan ekstraksi frase kunci dari layanan Bahasa untuk meringkas poin-poin utama.

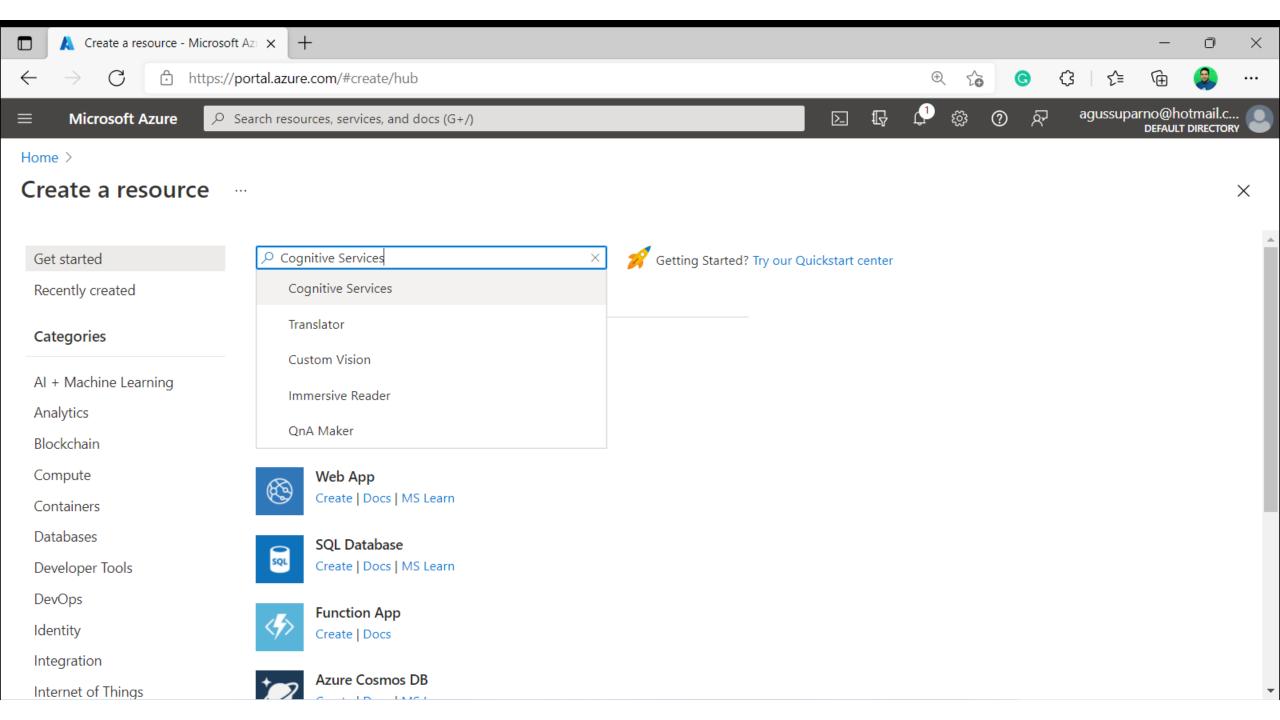
Entity recognition

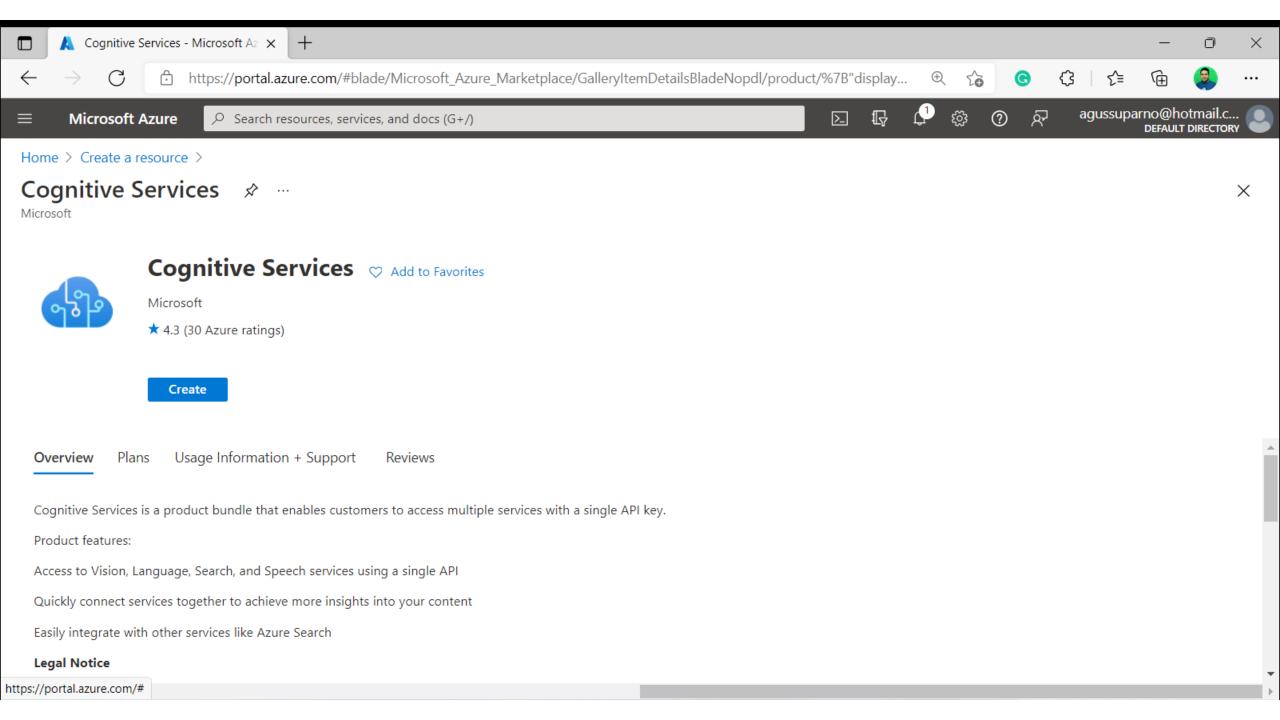
 Anda dapat menyediakan layanan "Language Service" dengan teks tidak terstruktur dan itu akan mengembalikan daftar entitas dalam teks yang dikenalinya. Layanan juga dapat menyediakan tautan ke informasi lebih lanjut tentang entitas tersebut di web. Entitas pada dasarnya adalah item dari jenis atau kategori tertentu; dan dalam beberapa kasus, subtipe, seperti yang ditunjukkan pada tabel berikut.

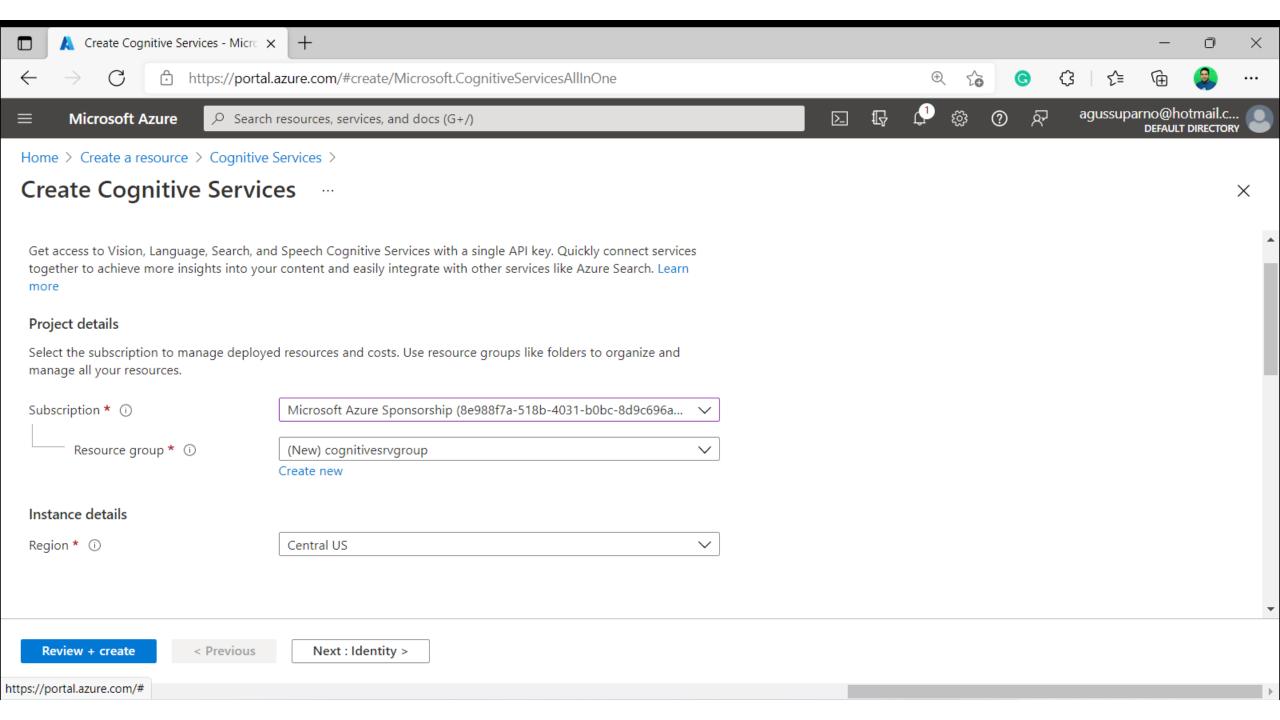
ENTITY RECOGNITION		
Туре	SubType	Example
Person		"Bill Gates", "John"
Location		"Paris", "New York"
Organization		"Microsoft"
Quantity	Number	"6" or "six"
Quantity	Percentage	"25%" or "fifty percent"
Quantity	Ordinal	"1st" or "first"
Quantity	Age	"90 day old" or "30 years old"
Quantity	Currency	"10.99"

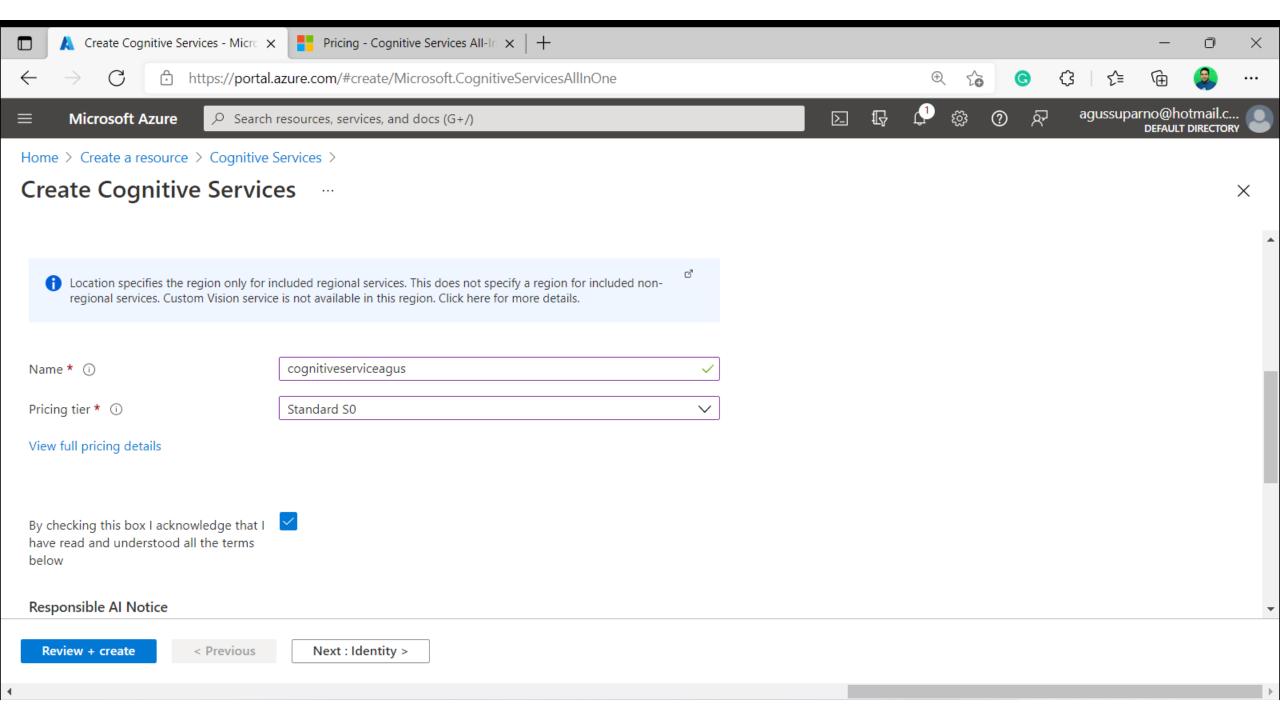
Demo Sentiment Analysis

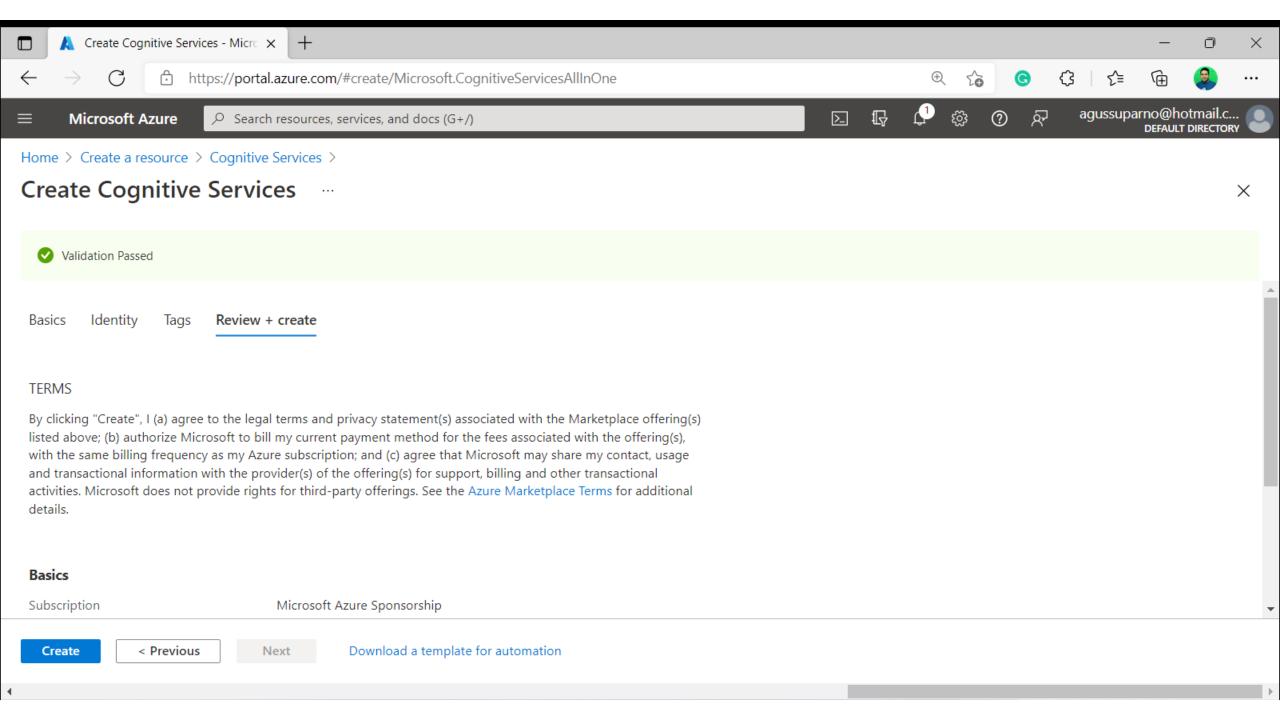


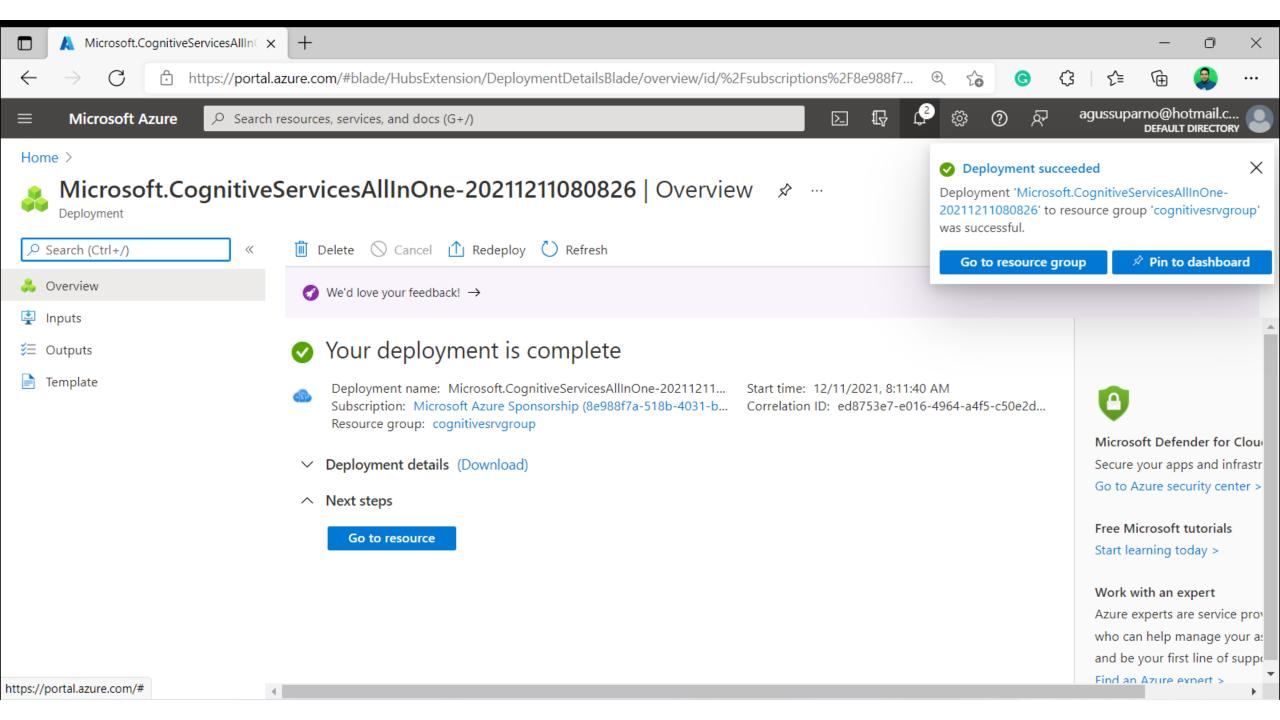


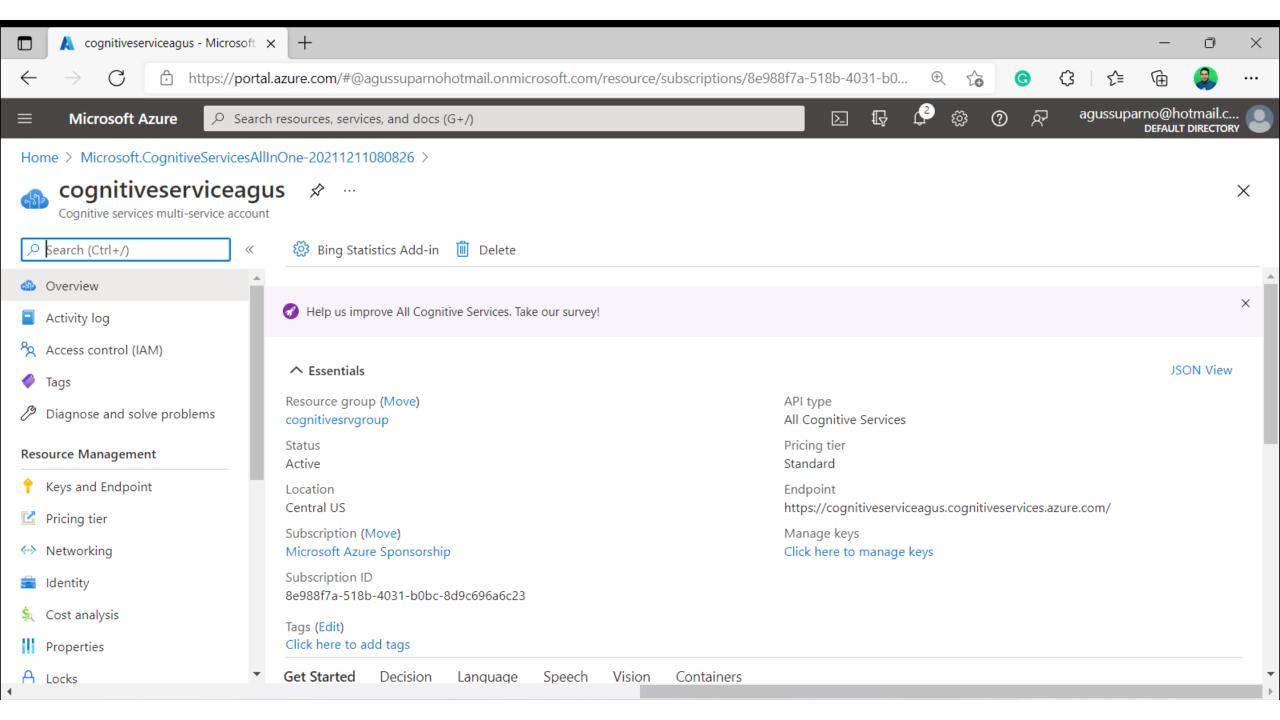


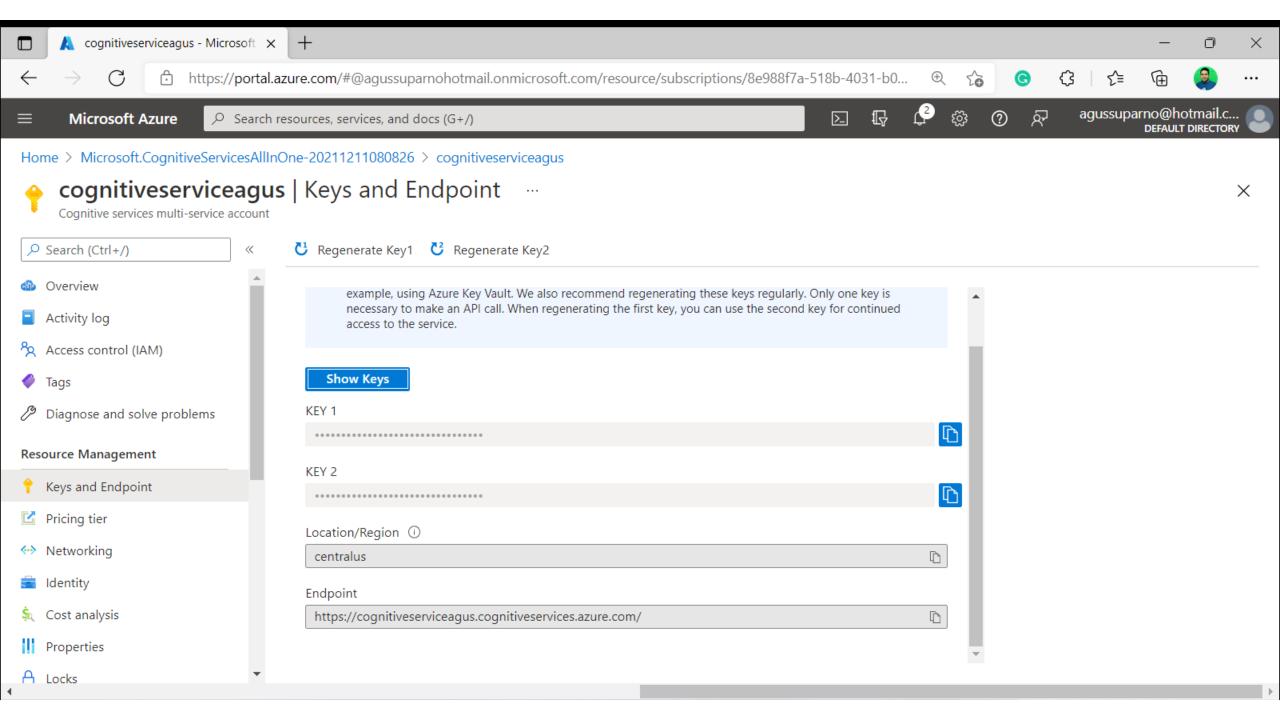


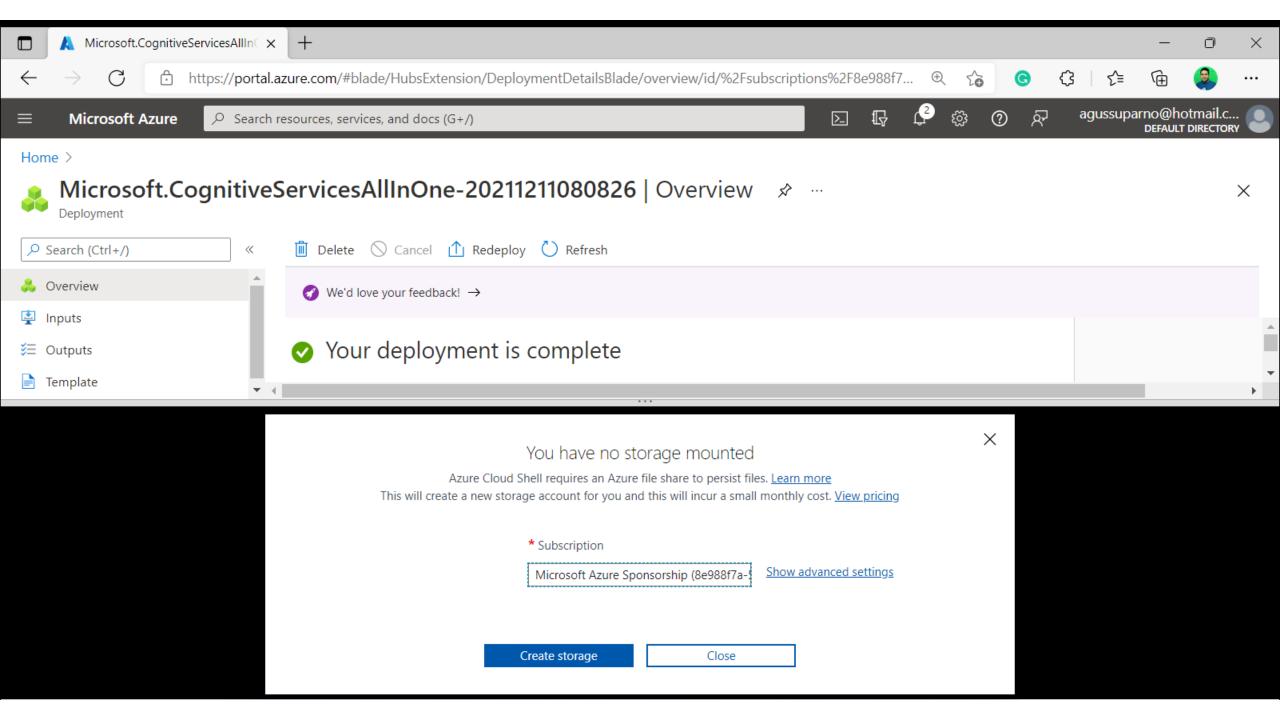


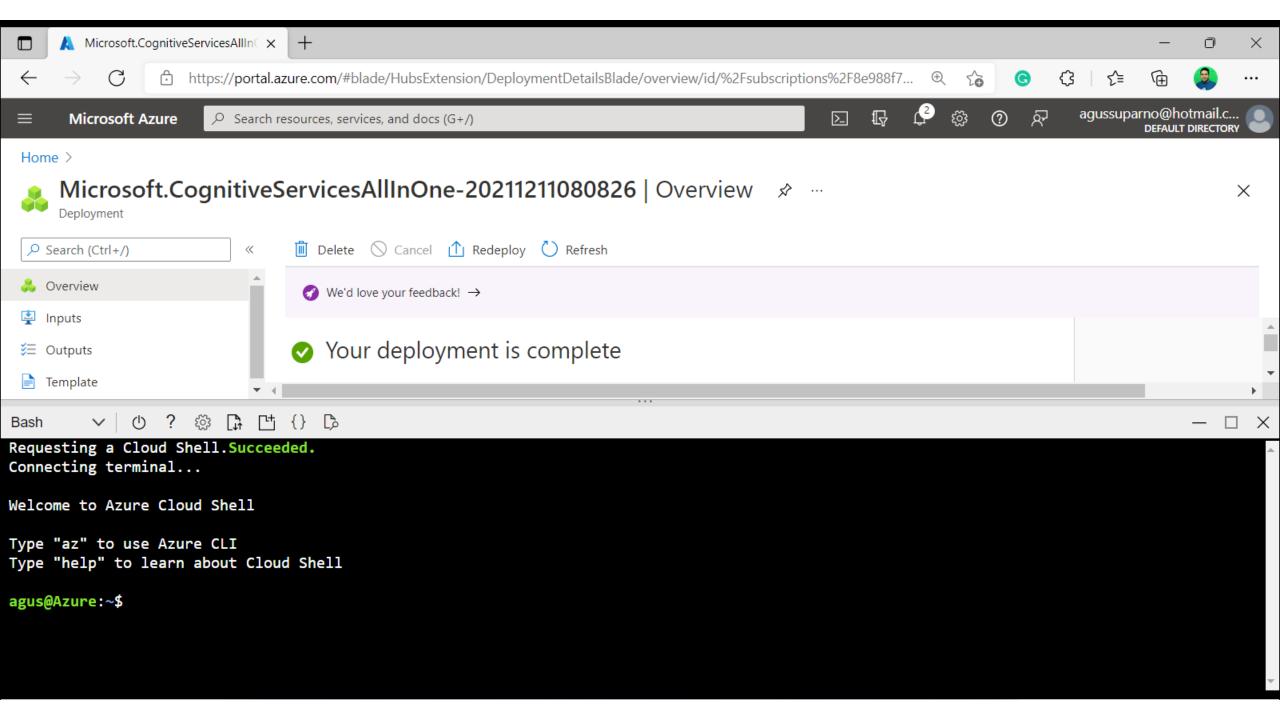






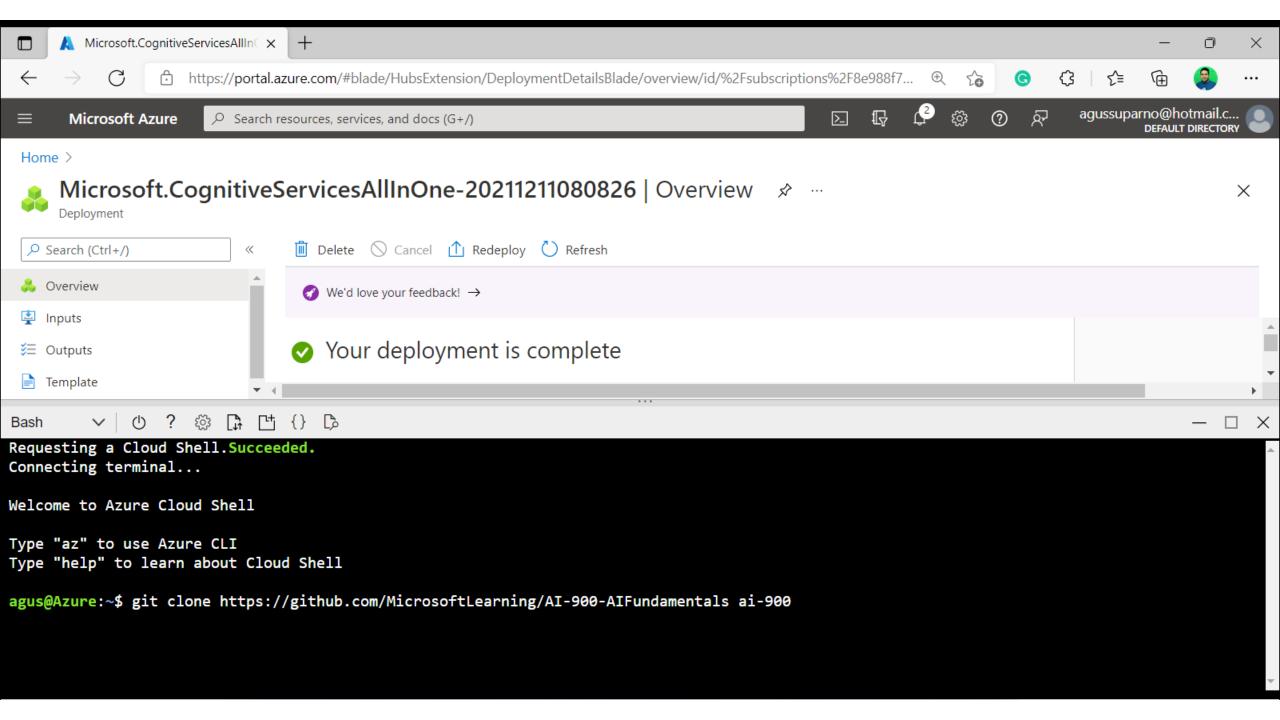


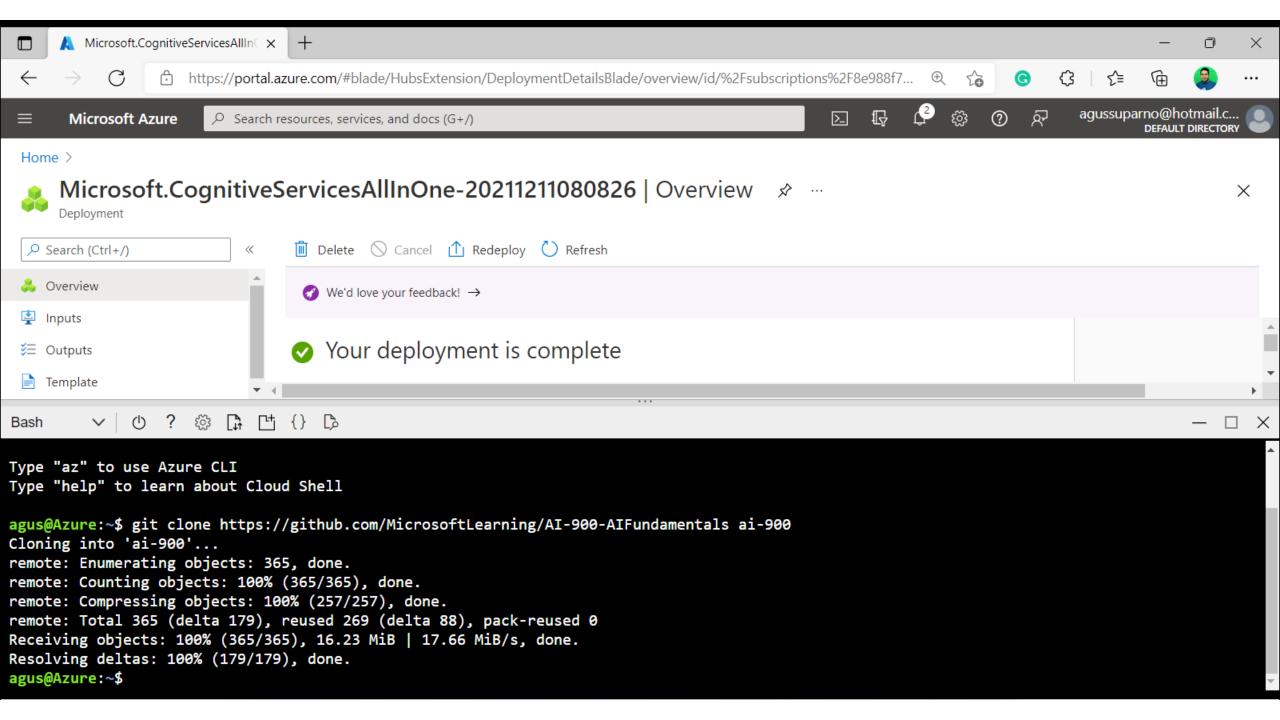


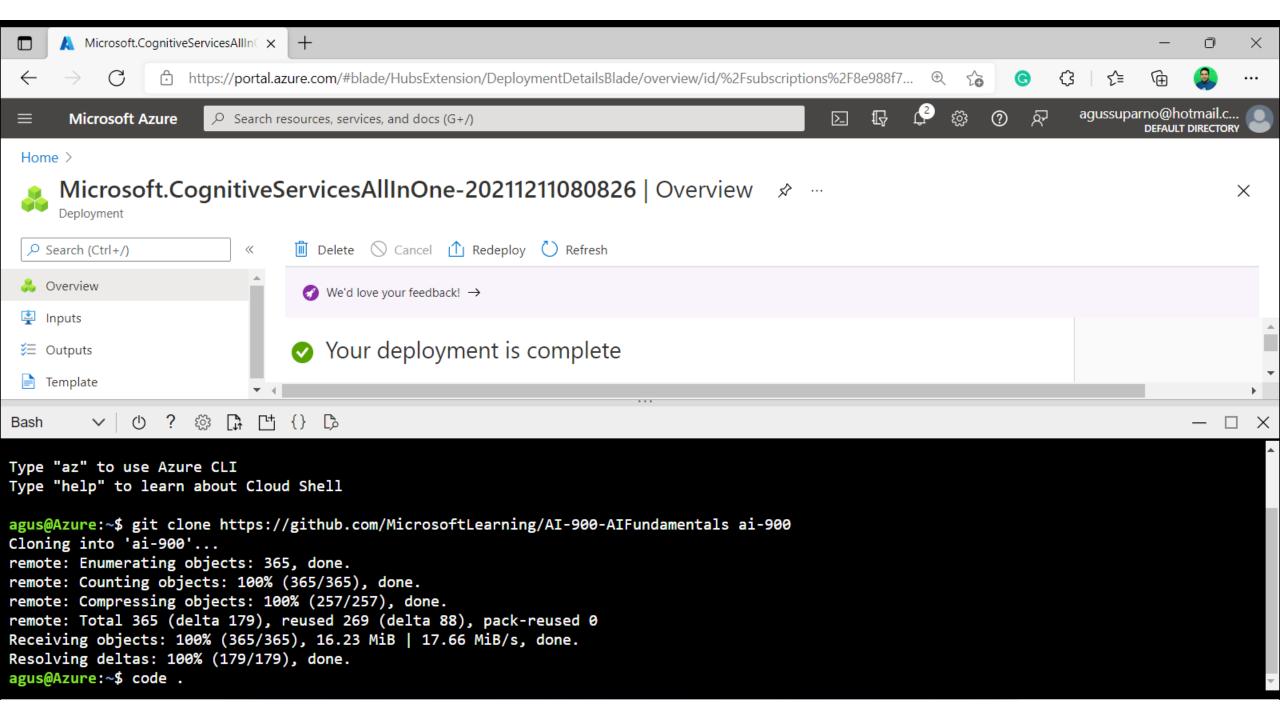


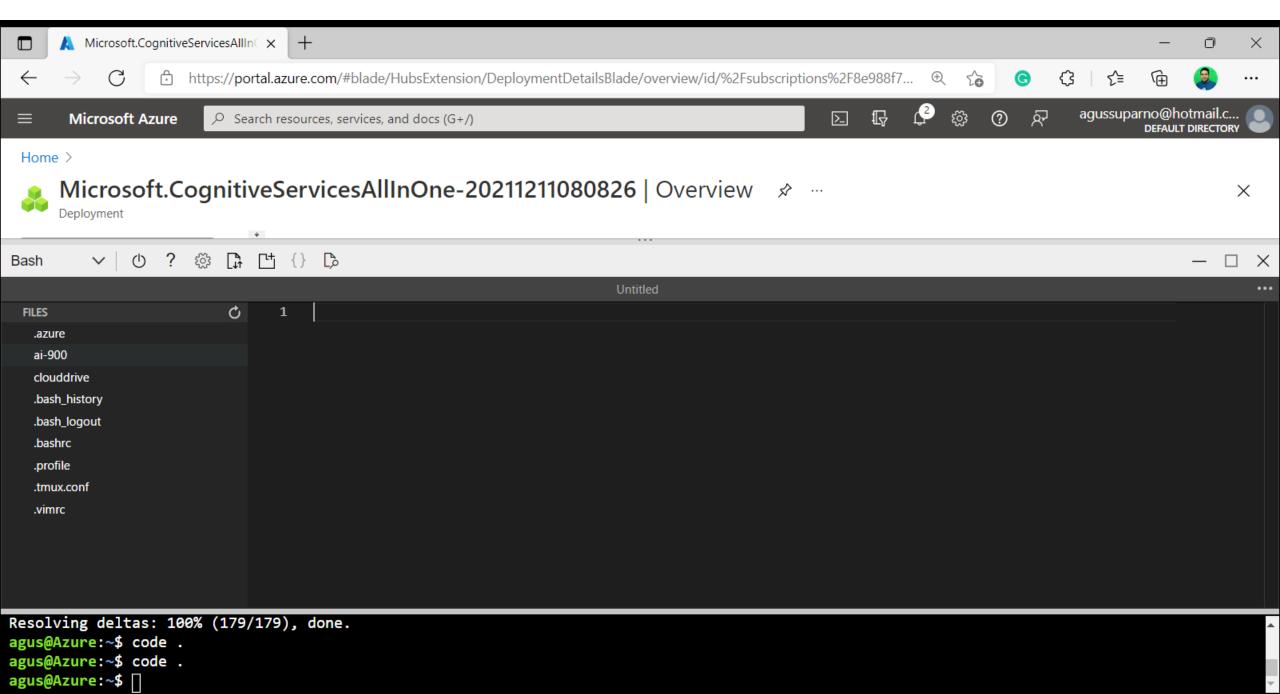
Clone code from AlFundamentals

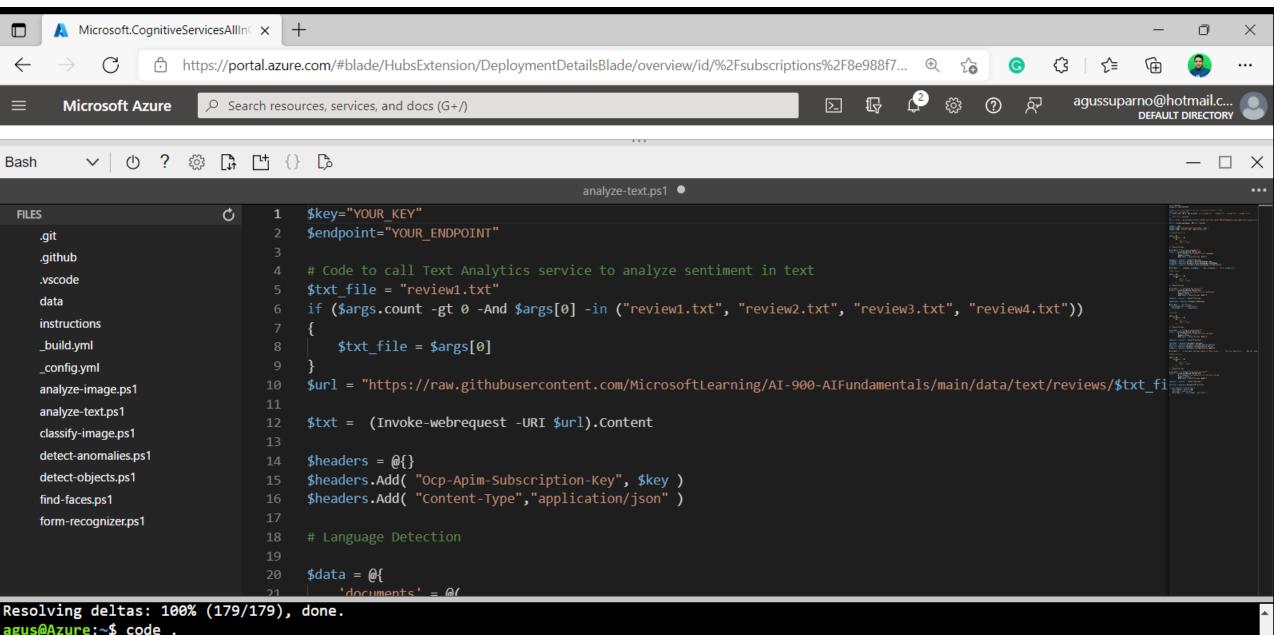
• git clone https://github.com/MicrosoftLearning/AI-900-AIFundamentals ai-900











agus@Azure:~\$ code .
agus@Azure:~\$ code .
agus@Azure:~\$ [

