Natural Language Processing Applications

Dr. Chetana Gavankar, Ph.D,
IIT Bombay-Monash University Australia
Chetana.gavankar@pilani.bits-pilani.ac.in

BITS Pilani
Pilani Campus

innovate achieve lead

BITS Pilani

Pilani Campus

Session 1

These slides are prepared by the instructor, with grateful acknowledgement of James Allen and many others who made their course materials freely available online.

Session Content

- Objective of course
- What will we learn in this course?
- Text books and Reference books
- Evaluation Plan
- Application areas of Natural Language Processing

Objective of course

CO1	Identify and describe algorithms for real life NLP Applications
CO2	Demonstrate understanding of algorithms by using different NLP tools
CO3	Apply NLP techniques in state of art applications like Machine Translation, Information Extraction including Named entity recognition and Relation extraction
CO4	Evaluate different approaches of implementing NLP applications along with ethical considerations

What you will learn in this course

- **Grammar and spellcheckers**
- **Question Answering and Conversational AI**
- **Knowledge Graph Applications**
- **Machine Translation**
  - Statistical
  - Neural
  - Indic Languages
- **Information Extraction**
  - Named Entity Recognition
  - Relation Extraction
  - Extracting Events and Time
- **Sentiment Analysis**
  - Sentiment Analysis Methods
  - Neural Networks for Sentiment Analysis

All above interesting and important real world applications will be discussed with case study and implementations in the respective modules.

Text books and Reference books

innovate achieve lead

T1	Speech and Language processing: An introduction to Natural Language Processing, Computational Linguistics and speech Recognition by Daniel Jurafsky and James H. Martin

R1	Manning and Schütze, Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA
R2	Neural Machine Translation by Philipp Koehn
R3	Knowledge Graphs Methodology, Tools and Selected Use Cases by Dieter Fensel , Umutcan Şimşek, Kevin Angele, Elwin Huaman , Elias Kärle , Oleksandra Panasiuk , Ioan Toma, Jürgen Umbrich, and Alexander Wahler, Springer 2019
R4	Natural Language Toolkit. Bird and Loper, and other developers. Available for free at: – http://www.nltk.org/

Evaluation Plan

Name	Weight
Quiz (best 2 out of 3)	10%
Assignment 1 and 2	30%
Mid-term Exam	30%
End Semester Exam	30%

What is Natural Language Processing?

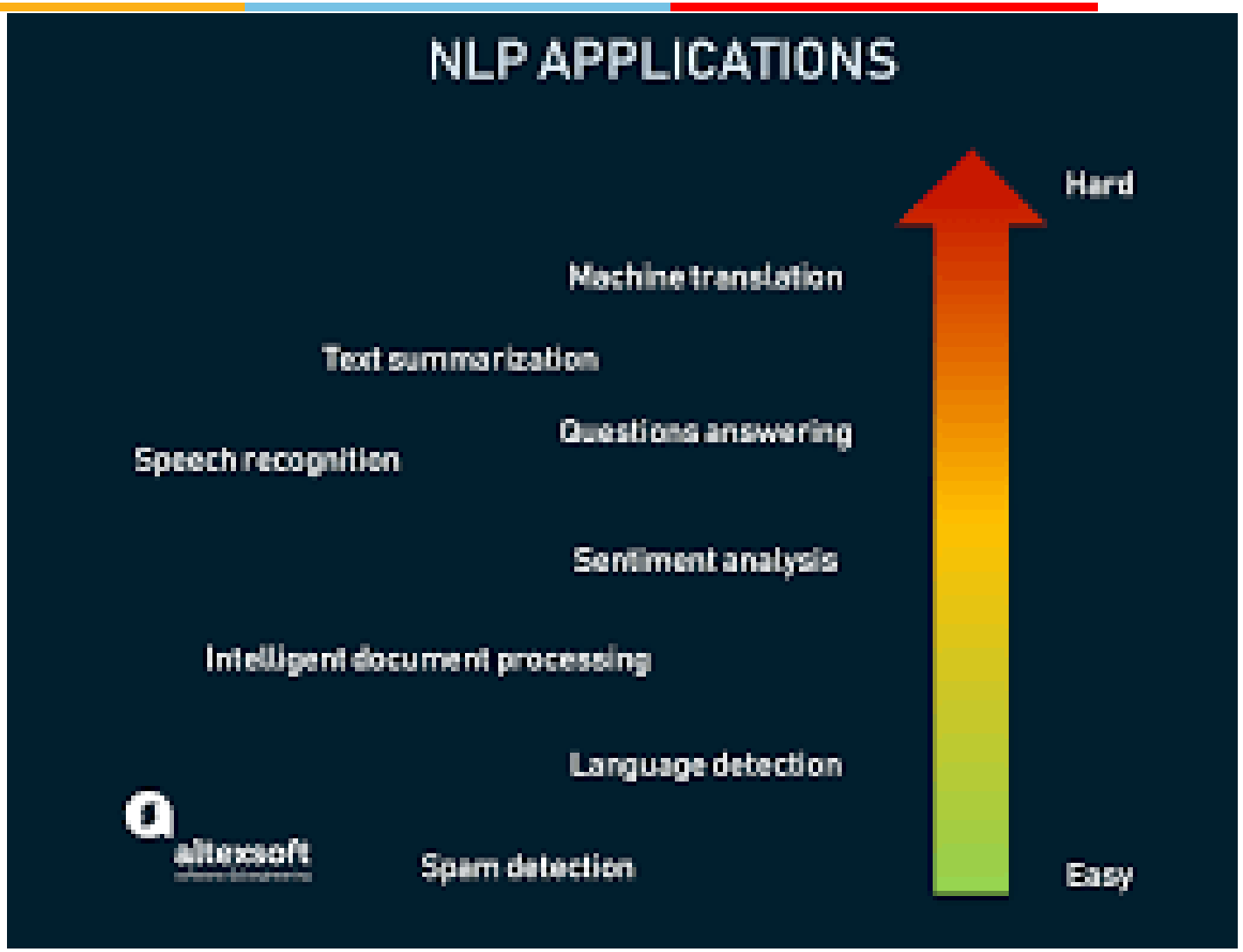
- Analyze, understand and generate human languages just like humans do
- Applying computational techniques to language domain
- To explain linguistic theories, to use the theories to build systems that can be of social use
- Make computers learn our language rather than we learn theirs.

NLP Tasks and Applications

It's a big world out there

And everyone uses language

NLP Applications and Difficulty Level



NLP Applications

- Question answering
  - Who is the first Taiwanese president?
- Text Categorization/Routing
  - e.g., customer e-mails.
- Text Mining
  - Find everything that can be done with NLP
- Machine (Assisted) Translation
- Language Teaching/Learning
  - Usage checking
- Spelling correction
  - Is that just dictionary lookup?

Application areas

- Text-to-Speech & Speech recognition
- Natural Language Dialogue Interfaces to Databases
- Information Extraction
- Document Image Analysis
- Automatic Summarization (https://pypi.org/project/sumy/)
- Text Proof-reading – Spelling & Grammar
- Machine Translation
- Fake News and Cyberbullying Detection
- Monitoring Social Media
- Plagiarism detection
- Question Answering System (https://haystack.deepset.ai/)
- Sentiment Analysis (https://komprehend.io/sentiment-analysis)

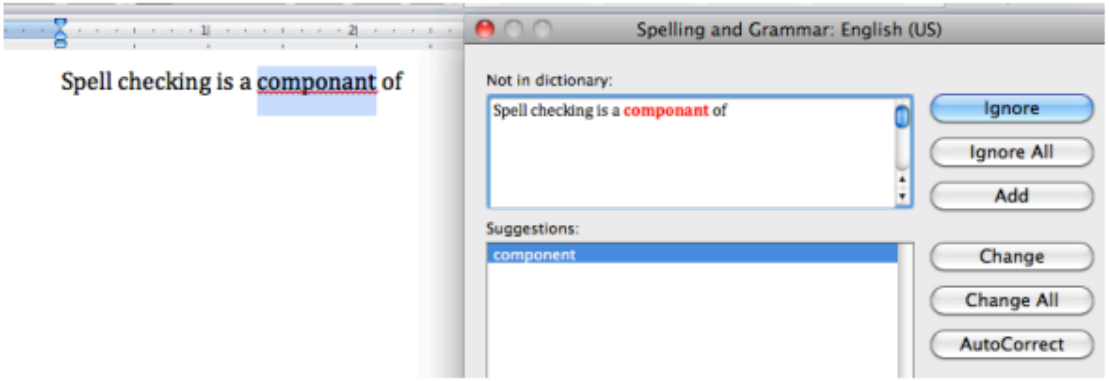
NLP Applications span across domains like Healthcare, Finance, Manufacturing and many more

Grammar and spellcheckers

Spell check

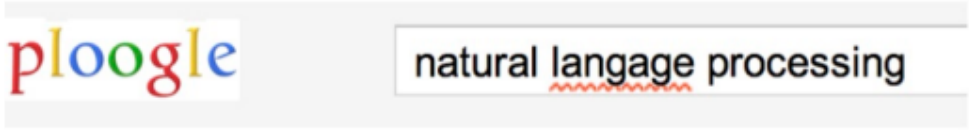
Applications for spelling correction

Word processing



Phones

Web search



2

Showing results for natural ***language*** processing
Search instead for natural langage processing

https://languagetool.org/

Spell check and Grammar check

Problem:

... and he fired presidential **aid/aide** Dick Morris after ...

⇒ *aid* or

⇒ *aide*

Training Data:

Spelling	Context
(1) aid	... and cut the foreign *aid/aide* budget in fiscal 1996 ...
" "	... they offered federal *aid/aide* for flood-ravaged states ...
(2) aide	... fired presidential *aid/aide* Dick Morris after ...
" "	... and said the chief *aid/aide* to Sen. Baker, Mr. John ...

Test Data:

Spelling	Context
???	... said the longtime *aid/aide* to the Mayor of St. ...
???	... will squander the *aid/aide* it receives from the ...

slide courtesy of D. Yarowsky

Case Study: AI-Powered Communication Enhancement

Challenges

- Traditional tools often miss contextual errors (e.g., *your* vs. *you're*, *affect* vs. *effect*).
- Writers need more than just error correction; they need help communicating effectively and appropriately for their audience and context.

SOLUTION

- **Hybrid Approach:** Combines rule-based checks, statistical NLP (analyzing patterns in vast text data), and advanced DL/Transformers/Agentic AI
- **Contextual Understanding:** AI models analyze the entire sentence and surrounding text to detect subtle errors and suggest contextually appropriate words.
- **Beyond Grammar:** Employs NLP to analyze tone (e.g., confident, formal, friendly), suggest improvements for clarity and conciseness, check for plagiarism, and offer full-sentence rewrites.
- **Continuous Learning:** Models are constantly updated based on user interactions and evolving language patterns

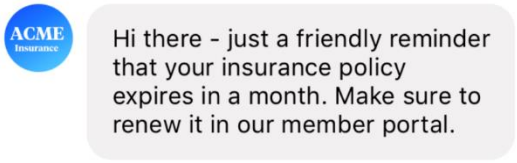
Question Answering and Conversational AI

Case Study: AI-powered Question Answering (QA) system

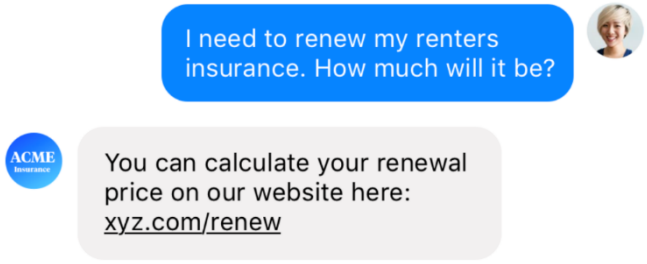
- 24/7 Demand, High volume, Long wait times
- **Training:** The AI was trained to accurately handle the top 100 most frequent customer questions.
- **Integration:** The QA chatbot was launched, acting as the *first point of contact* for all support inquiries.
- **Handoff):** A seamless "handoff" protocol was created, allowing the bot to escalate complex or sensitive issues to a live human agent, along with the full chat history.

AI Assistants

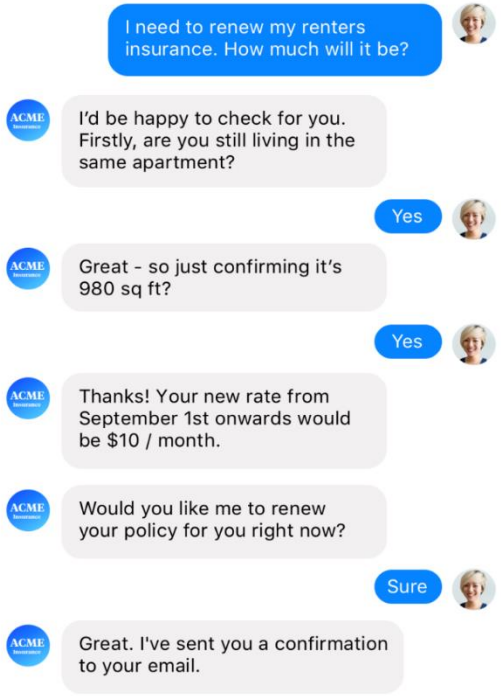
Notification Assistant



FAQ Assistant



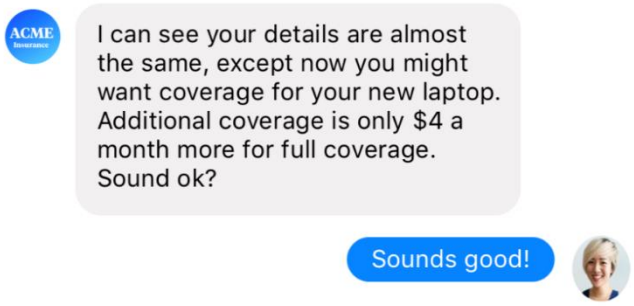
Contextual Assistant



AI Assistants

Personalized Assistant

- Assistant knows you much more in detail
- Quickly checks a few final things before giving you a quote tailored to your actual situation.



Autonomous Organization of Assistants

- Group of AI assistants that know every customer personally
- Eventually run large parts of company operations—from lead generation over marketing, sales, HR, or finance

Case Study: Automating Customer Support with NLP Chatbots

innovate achieve lead

- **Background**
- HDFC Bank, one of India's largest private sector banks, was facing increasing customer service demands. Traditional support channels (phone/email) were overburdened, resulting in delayed responses and reduced customer satisfaction.
- **Problem**
- High customer query volume (especially repetitive questions)
- Long response times
- High operational costs for customer support
- **Solution: EVA – HDFC Bank's AI Chatbot**

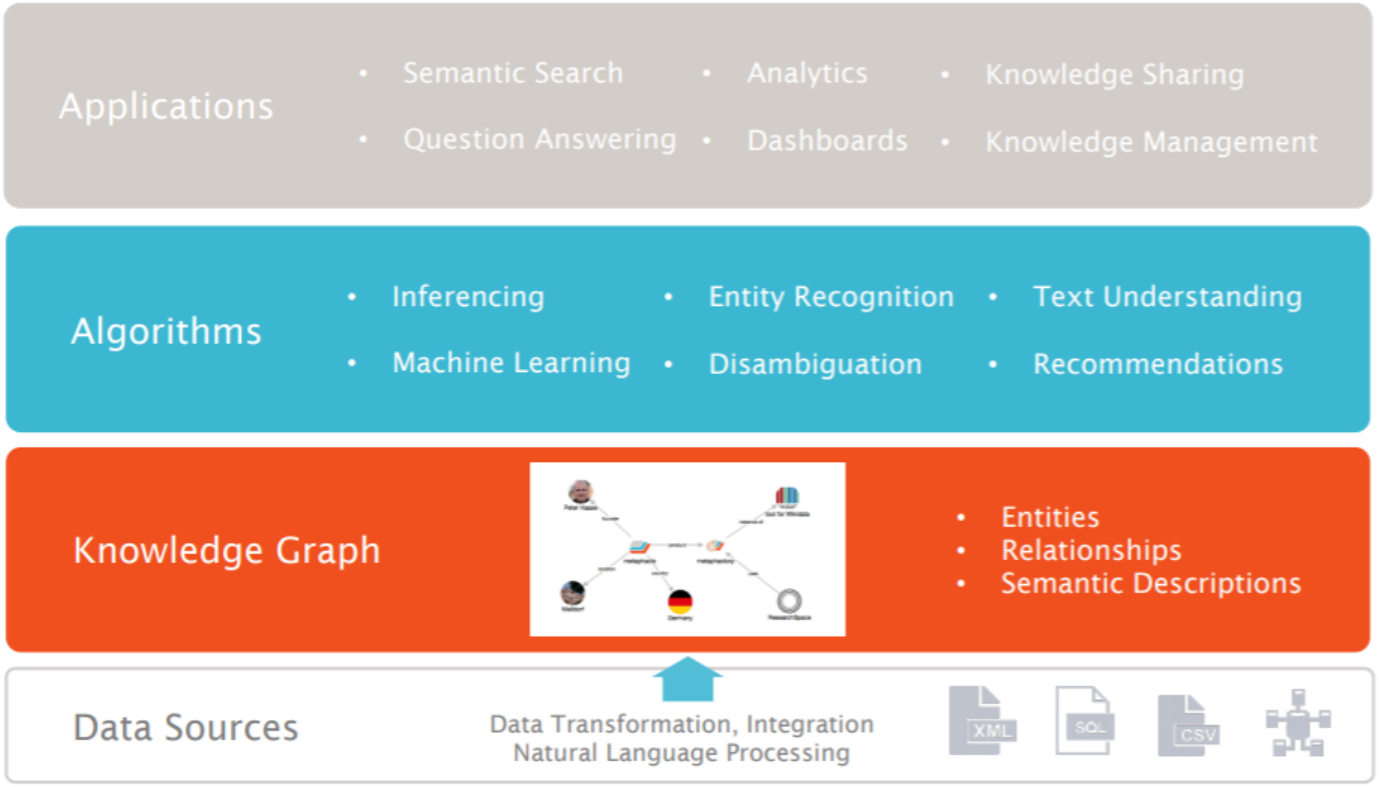
Case Study: Automating Customer Support with NLP Chatbots

- **Key NLP Features Used**
  - **Intent recognition**: Classifies queries into categories (e.g., account balance, card issues)
  - **Named Entity Recognition (NER)**: Identifies dates, transaction amounts, account types
  - **Context management**: Maintains conversational state for multi-turn dialogues
  - **Multilingual support**: Handles queries in English and Hindi
- **Outcomes**
  - Handled over **2.7 million queries** in the first year
  - Resolved **85% of queries** without human intervention
  - **Reduced average query response time** to less than 0.4 seconds
  - Improved **customer satisfaction** and reduced cost-per-query
- **Challenges Faced**
  - Training the model on diverse and noisy real-world customer data
  - Maintaining user trust (transparency and security)
  - Escalating complex queries to human agents without losing context

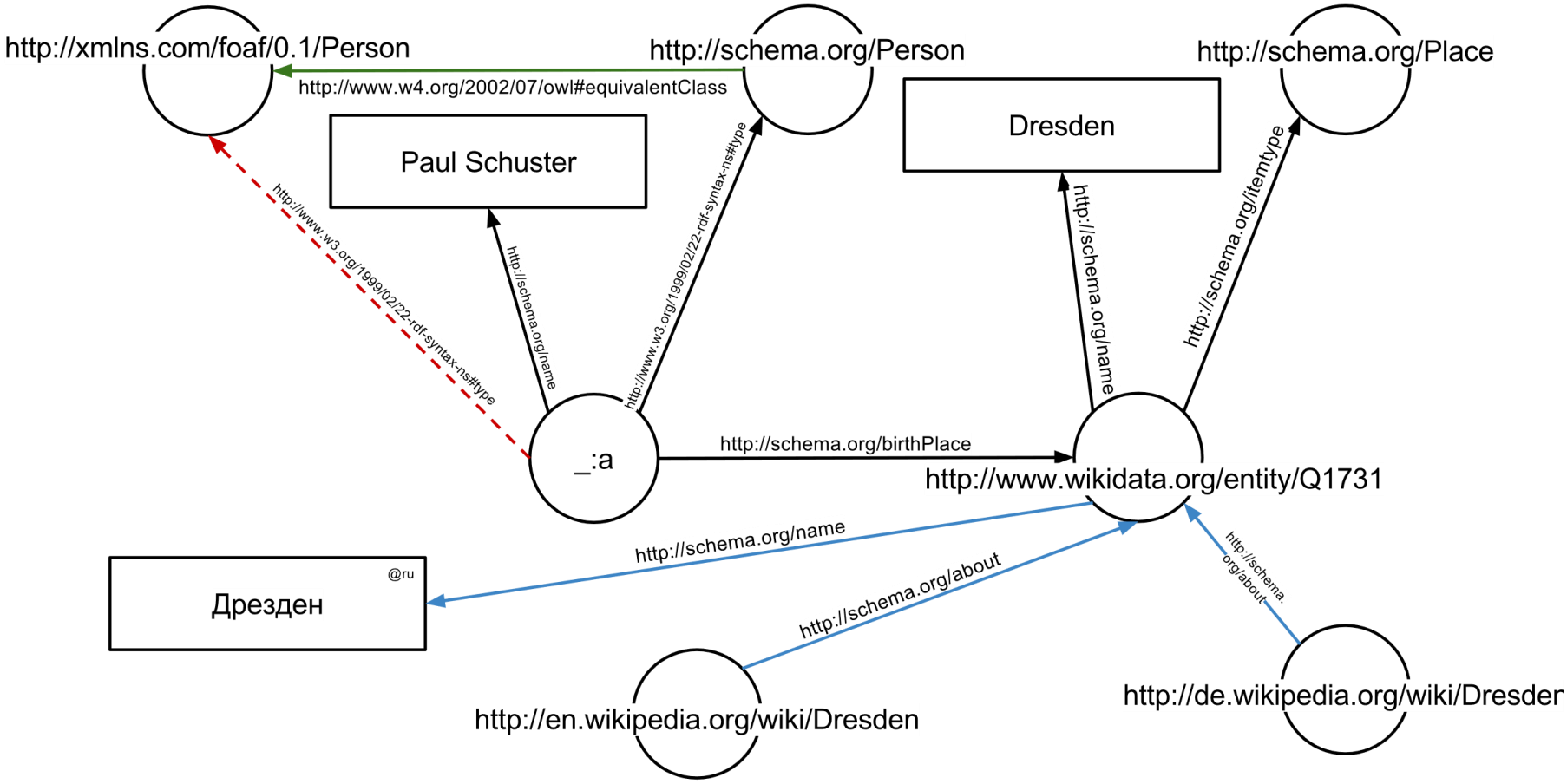
Knowledge Graph Applications

Knowledge Graph

- Knowledge Graph Applications



The Semantic Web



from en.wikipedia.org

Case study – Knowledge Graph Applications in Agentic AI

Problem: Large tech company, faced challenges with internal information access. Key data about projects, teams, expertise, and documentation was scattered across various internal wikis, project management tools (like Jira), employee directories, and shared drives.

Employees spent significant time searching for answers to complex questions like:

- "Who worked on the authentication module for Project X and what were the key technical documents?"
- "Which active projects are using Python and are related to our client Y?"
- "Find the design specification document for the UI redesign led by A's team."

Enterprise Knowledge Graph (EKG) Construction

Data Sources: Internal Wiki, Jira, HR Database, Code Repositories (metadata), Document Management System.

Entities: Employee, Project, Team, Document, Skill, Client, CodeModule.

Relationships: WORKS_ON (Employee -> Project), MEMBER_OF (Employee -> Team), AUTHORED (Employee -> Document), RELATED_TO (Project -> Document), HAS_SKILL (Employee -> Skill), SERVES_CLIENT (Project -> Client), IMPLEMENTS (CodeModule -> Project).

Population: NLP techniques (NER, Relation Extraction) were used to extract entities and relationships from unstructured text (wikis, documents). Structured data was mapped directly (HR DB, Jira)

Agentic AI System ("Ask Me")

Core: An LLM (like GPT-4 or a fine-tuned open-source model)

- **Tools:**
  - query_EKG: Takes a natural language query part and translates it into a structured query (e.g., Cypher for Neo4j) to fetch data from the EKG.
  - document_search: Performs keyword/semantic search over the document system for full-text retrieval when needed.
- **Capabilities:** Planning, Tool Use, Reasoning, Synthesis.

Example of Enterprise KG application

User Query: "Who led the backend development for Project X and what recent documents did they author?"

Agent's Plan:

1.Identify entities: "Project X" (Project), "backend development" (implicit Skill/Role), "documents" (Document).

2.Find the team/person associated with "Project X" and "backend development lead" role via EKG. -> Use query_EKG.

3.Once the person is identified (e.g., “A"), find recent documents AUTHORED by “A" via EKG. -> Use query_EKG.

4.Synthesize the results into a natural language answer.

Execution:

•Agent calls query_EKG ("Find lead backend developer for Project X").
EKG returns "John Smith".

•Agent calls query_EKG ("Find recent documents authored by John Smith"). EKG returns a list: ["API_Design_v2.pdf", "Deployment_Strategy.docx"].

Agent's Response: “A led the backend development for Project X. Recent documents he authored include 'API_Design_v2.pdf' and Deployment_Strategy.docx'."

Machine Translation

Machine Translation

English to Hindi http://anglahindi.iitk.ac.in

Simple Sentences.
sarala vaakya .

Welcome to London.
landana men aapakaa svaagata hai.

There are some cases which are still pending.

NLP applications is an interesting course in MTech AIML of BITS WILP

एनएलपी एप्लिकेशन बिट्स WILP के एमटेक एआईएमएल में एक दिलचस्प कोर्स है

Machine Translation - Challenges

Accent Restoration in Spanish & French

Problem:

Input: ... deja travaille cote a cote ...

⇓

Output: ... déjà travaillé côte à côte ...

Examples:

... appeler l'autre **cote** de l'atlantique ...

⇒ *côté* (meaning side) or

⇒ *côte* (meaning coast)

... une famille des **pecheurs** ...

⇒ *pêcheurs* (meaning fishermen) or

⇒ *pécheurs* (meaning sinners)

slide courtesy of D. Yarowsky

Machine Translation

Accent Restoration in Spanish & French

Training Data:

Pattern	Context
(1) côté	... du laisser de *cote* faute de temps ...
” ”	... appeler l’ autre *cote* de l’ atlantique ...
” ”	... passe de notre *cote* de la frontiere ...
(2) côte	... vivre sur notre *cote* ouest toujours ...
” ”	... creer sur la *cote* du labrador des ...
” ”	travaillaient cote a *cote* , ils avaient ...

Test Data:

Pattern	Context
???	... passe de notre *cote* de la frontiere ...
???	... creer sur la *cote* du labrador des ...

slide courtesy of D. Yarowsky

Machine Translation - Challenges

Capitalization

Problem:

... FRIED CHICKEN, **TURKEY** SANDWICHES AND FROZEN ...

⇒ *turkey* (the *bird*) or

⇒ *Turkey* (the *country*)

Training Data:

Capitalization	Context
(1) turkey	... OF FRIED CHICKEN , TURKEY SANDWICHES AND FROZEN ...
" "	... NTS A POUND , WHILE TURKEY PRICES ROSE 1.2 CENTS ...
" "	... PLAY , REAL GRADE-A TURKEY , WHICH ONLY A PRICE ...
(2) Turkey	... INUNDATED EASTERN TURKEY AFTER THE EARLIER ...
" "	... FEELINGS TOWARD TURKEY SURFACED WHEN GREECE ...
" "	... THE CONTRACT WITH TURKEY WILL PROVIDE OPPORTU...

Test Data:

Capitalization	Context
???	... NECK LIKE THAT OF A TURKEY ON A CHOPPING BLOCK ...
???	... PROBLEM IS THAT TURKEY IS NOT A EUROPEAN ...

slide courtesy of D. Yarowsky

Machine Translation - Challenges

Word sense disambiguation

Problem:

... He wrote the last **sentence** two years later ...

⇒ *sentencia* (legal sentence) or

⇒ *frase* (grammatical sentence)

Training Data:

Translation	Context
(1) sentencia	... for a maximum *sentence* for a young offender ...
" "	... of the minimum *sentence* of seven years in jail ...
" "	... were under the *sentence* of death at that time ...
(2) frase	... read the second *sentence* because it is just as ...
" "	... The next *sentence* is a very important ...
" "	... It is the second *sentence* which I think is at ...

Test Data:

Translation	Context
???	... cannot criticize a *sentence* handed down by ...
???	... listen to this *sentence* uttered by a former ...

Machine Translation - Challenges

Text to speech

Problem:

... slightly elevated *lead* levels ...

⇒ *lɛd* (as in *lead mine*) or

⇒ *li:d* (as in *lead role*)

Training Data:

Pronunciation	Context
(1) lɛd	... it monitors the *lead* levels in drinking ...
" "	... conference on *lead* poisoning in ...
" "	... strontium and *lead* isotope zonation ...
(2) li:d	... maintained their *lead* Thursday over ...
" "	... to Boston and *lead* singer for Purple ...
" "	... Bush a 17-point *lead* in Texas , only 3 ...

Test Data:

Pronunciation	Context
???	... median blood *lead* concentration was ..
???	... his double-digit *lead* nationwide . The ...

slide courtesy of D. Yarowsky

Case Study: Scaling Global Communications

- **Custom Engines:** The system was trained on specific content to create custom models that understood its unique terminology and branding.

- **Centralized Platform:** All translation requests were funneled through the MT platform, which was integrated into existing content systems.

- **MT + Post-Editing:** The platform was used to instantly translate content. For high-visibility materials, human translators would then "post-edit" the AI's output, rather than translating from scratch.

Information Extraction

Information Extraction

As a task:

Filling slots in a database from sub-segments of text.

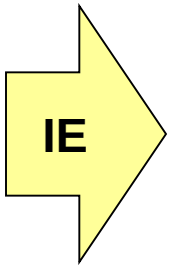
October 14, 2002, 4:00 a.m. PT

For years, Microsoft Corporation CEO Bill Gates railed against the economic philosophy of open-source software with Orwellian fervor, denouncing its communal licensing as a "cancer" that stifled technological innovation.

Today, Microsoft claims to "love" the open-source concept, by which software code is made public to encourage improvement and development by outside programmers. Gates himself says Microsoft will gladly disclose its crown jewels--the coveted code behind the Windows operating system--to select customers.

"We can be open source. We love the concept of shared source," said Bill Veghte, a Microsoft VP. "That's a super-important shift for us in terms of code access."

Richard Stallman, founder of the Free Software Foundation, countered saying…



NAME	TITLE	ORGANIZATION
Bill Gates	CEO	Microsoft
Bill Veghte	VP	Microsoft
Richard Stallman	founder	Free Soft..

Slide from Chris Brew, adapted from slide by William Cohen

Phrase Types to Identify for IE

Closed set

U.S. states

He was born in Alabama…

The big Wyoming sky…

Regular set

U.S. phone numbers

Phone: (413) 545-1323

The CALD main office can be reached at 412-268-1299

Complex pattern

U.S. postal addresses

University of Arkansas
P.O. Box 140
Hope, AR 71802

Headquarters:
1128 Main Street, 4th Floor
Cincinnati, Ohio 45210

Ambiguous patterns, needing context and many sources of evidence

Person names

…was among the six houses sold by Hope Feldman that year.

Pawel Opalinski, Software Engineer at WhizBang Labs.

Slide from Chris Brew, adapted from slide by William Cohen

Deeper Information Extraction

1. Coreference resolution (within a document)
2. Entity linking (across documents)
3. Event extraction and linking
4. Knowledge base population (KBP)

Named Entity Recognition

CHICAGO (AP) — Citing high fuel prices, United Airlines said Friday it has increased fares by $6 per round trip on flights to some cities also served by lower-cost carriers. American Airlines, a unit AMR, immediately matched the move, spokesman Tim Wagner said. United, a unit of UAL, said the increase took effect Thursday night and applies to most routes where it competes against discount carriers, such as Chicago to Dallas and Atlanta and Denver to San Francisco, Los Angeles and New York.

NE Types

Type	Tag	Sample Categories
People	PER	Individuals, fictional characters, small groups
Organization	ORG	Companies, agencies, political parties, religious groups, sports teams
Location	LOC	Physical extents, mountains, lakes, seas
Geo-Political Entity	GPE	Countries, states, provinces, counties
Facility	FAC	Bridges, buildings, airports
Vehicles	VEH	Planes, trains, and automobiles

Type	Example
People	*Turing* is often considered to be the father of modern computer science.
Organization	The *IPCC* said it is likely that future tropical cyclones will become more intense.
Location	The *Mt. Sanitas* loop hike begins at the base of *Sunshine Canyon*.
Geo-Political Entity	*Palo Alto* is looking at raising the fees for parking in the University Avenue district.
Facility	Drivers were advised to consider either the *Tappan Zee Bridge* or the *Lincoln Tunnel*.
Vehicles	The updated *Mini Cooper* retains its charm and agility.

Slide from Jim Martin

Identifying phrases (chunking)

- Phrases that are useful for information extraction:
  - Named entities
    - As on previous slides
  - Relationship phrases
    - “said”, “according to”, …
    - “was born in”, “hails from”, …
    - “bought”, “hopes to acquire”, “formed a joint agreement with”, …
  - Simple syntactic chunks (e.g., non-recursive NPs)
    - “Syntactic chunking” sometimes done before (or instead of) parsing
    - Also, “segmentation”: divide Chinese text into words (no spaces)
- So, how do we learn to mark phrases?
  - Earlier, we built an FST to mark dates by inserting brackets
  - But, it’s common to set this up as a tagging problem …

Reduce to a tagging problem …

- The IOB encoding (Ramshaw & Marcus 1995):
  - B_X = "beginning" (first word of an X)
  - I_X = "inside" (non-first word of an X)
  - O = "outside" (not in any phrase)
  - Does not allow overlapping or recursive phrases

…United Airlines said Friday it has increased …

B_ORG I_ORG O O O O O

… the move , spokesman Tim Wagner said …

O O O O B_PER I_PER O

What if this were tagged as B_ORG instead?

Slide adapted from Chris Brew

Example applications for IE

- Classified ads
- Restaurant reviews
- Bibliographic citations
- Appointment emails
- Legal opinions
- Papers describing clinical medical studies
- …
- Adding facts to the semantic web

Sentiment Analysis

Sentiment classification

What features of the text could help predict # of stars?
(e.g., using a log-linear model) How to identify more?
Are the features hard to compute? (syntax? sarcasm?)

★☆☆☆☆ **An extremely versatile machine!**, November 22, 2006

By **Dr. Nickolas E. Jorgensen "njorgens3"**

This review is from: Cuisinart DGB-600BC Grind & Brew, Brushed Chrome (Kitchen)

This coffee-maker does so much! It makes weak, watery coffee! It grinds beans if you want it to! It inexplicably floods the entire counter with half-brewed coffee when you aren't looking! Perhaps it could be used to irrigate crops... It is time-consuming to clean, but in fairness I should also point out that the stainless-steel thermal carafe is a durable item that has withstood being hurled onto the floor in rage several times. And if all these features weren't enough, it's pretty expensive too. If faced with the choice between having a car door repeatedly slamming into my genitalia and buying this coffee-maker, I'd unhesitatingly choose the Cuisinart! The coffee would be lousy, but at least I could still have children...

Other text categorization tasks

- Is it spam? (see features)
- What grade, as an answer to this essay question?
- Is it interesting to this user?
  - News filtering; helpdesk routing
- Is it interesting to this NLP program?
  - Skill classification for a digital assistant!
  - If it's Spanish, translate it from Spanish
  - If it's subjective, run the sentiment classifier
  - If it's an appointment, run information extraction
- Where should it be filed?
  - Which mail folder? (work, friends, junk, urgent ...)
  - Yahoo! / Open Directory / digital libraries

Case Study: AI-powered sentiment analysis platform

Data Aggregation:

•The tool pulled all reviews into a single dashboard.

Aspect-Based Sentiment:

•The AI didn't just label a review "Positive" or "Negative.“

• It identified specific topics (aspects) and the sentiment for each.

Real-time Dashboard:

- Operations managers could instantly filter feedback by property, region, date, and topic.

Evaluation of NLP Applications

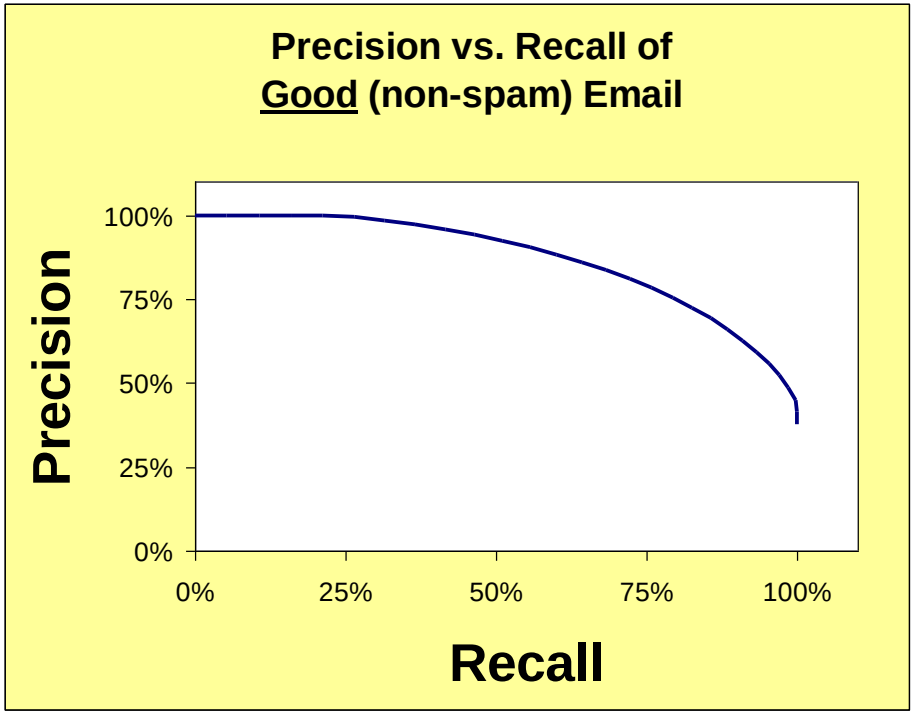
Measuring Performance

- Classification accuracy: What % of messages were classified correctly?
- **Is this what we care about?**

	Overall accuracy	Accuracy on spam	Accuracy on gen
System 1	95%	99.99%	90%
System 2	95%	90%	99.99%

- Which system do you prefer?

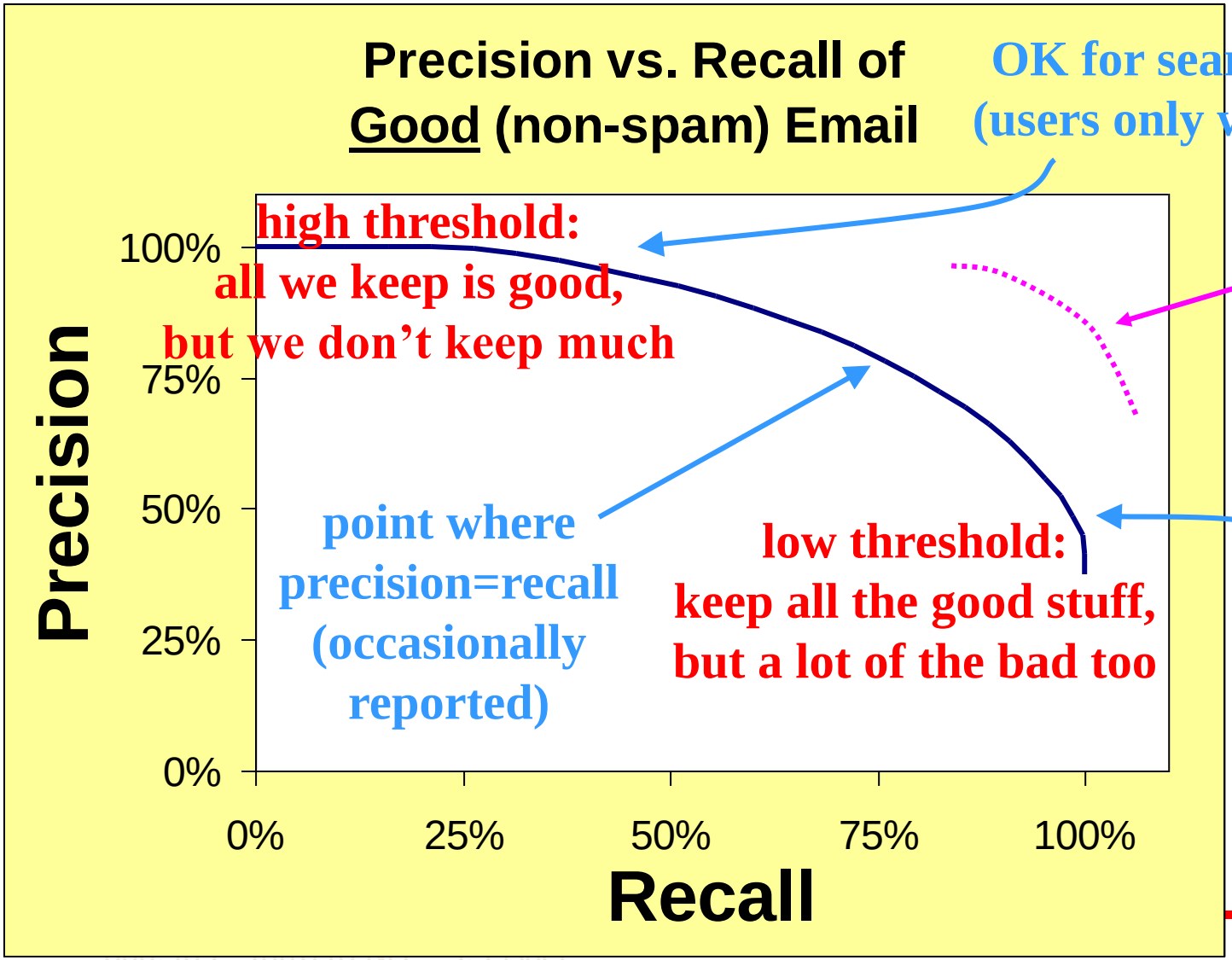
Measuring Performance



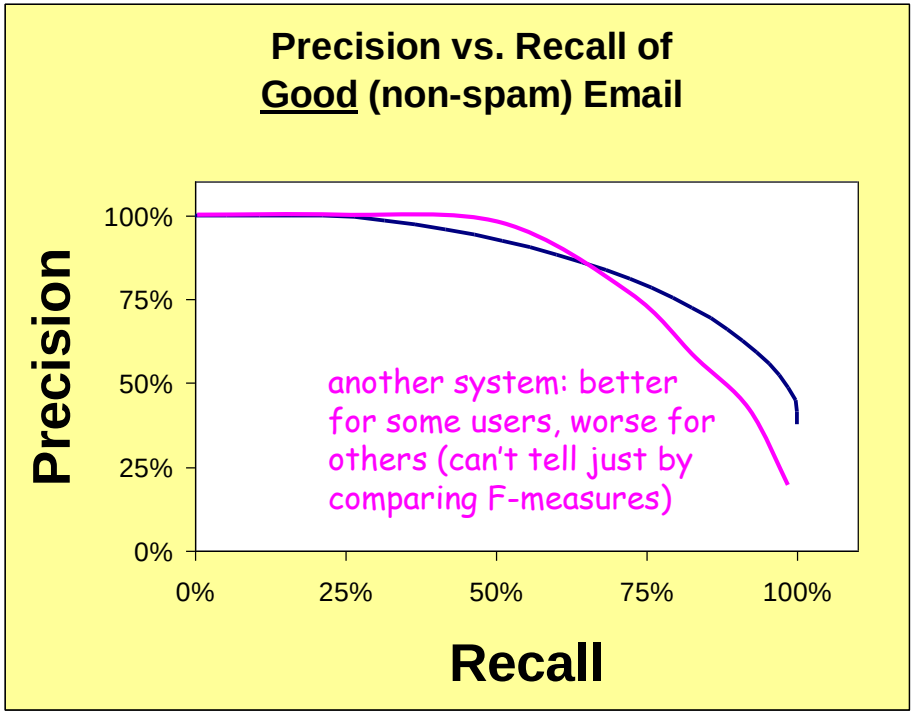
- Precision = $\frac{\text{good messages kept}}{\text{all messages kept}}$
- Recall = $\frac{\text{good messages kept}}{\text{all good messages}}$

Move from high precision to high recall by deleting fewer messages (delete only if spamminess > high threshold)

Measuring Performance



Measuring Performance

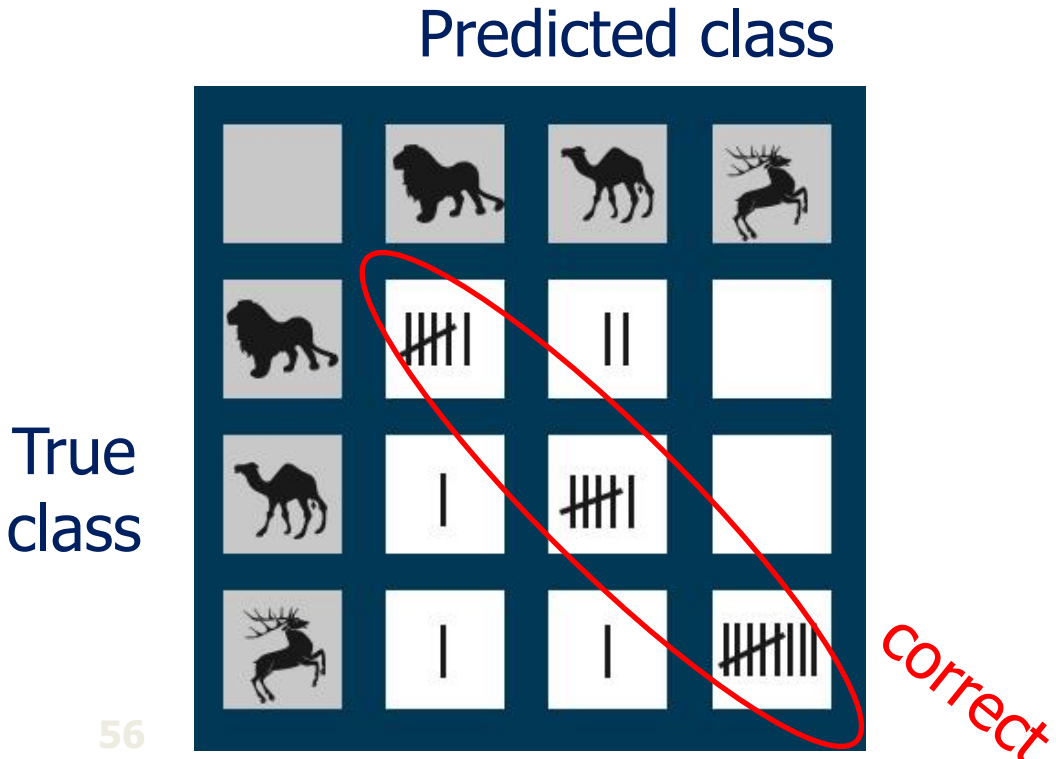


- Precision = $\frac{\text{good messages kept}}{\text{all messages kept}}$
- Recall = $\frac{\text{good messages kept}}{\text{all good messages}}$
- F-measure = $\left(\frac{\text{precision}^{-1} + \text{recall}^{-1}}{2}\right)^{-1}$

- Move from high precision to high recall by deleting fewer messages (raise threshold)
- Conventional to tune system and threshold to optimize F-measure on dev data
- But it's more informative to report the whole curve
  - Since in real life, the user should be able to pick a tradeoff point they like

More than 2 classes

- Report F-measure for <u>each</u> class
- Show a confusion matrix



Generating new text

1. Speech recognition (transcribe as text)
2. Machine translation
3. Text generation from semantics
4. Inflect, analyze, pronounce, or transliterate words
5. Single- or multi-doc summarization

Multimodal interfaces or modeling

1. Sign languages
2. Speech + gestures
3. Images + captions
4. Brain recordings, human reaction times

Some Big Questions

- Neural nets are fluent at generating text, but do they really represent and reason about the world the text describes? Are their answers consistent? Can they explain them?
- How can models learn effectively through interaction with the world or with human teachers?
- What kinds of linguistic biases should we build in, and how? Huge Transformer LMs with enormous training sets work well, but can we find architectures that generalize like humans from much smaller datasets? (Or is that just pretraining + few-shot or fine-tuning?)

The NLP Research Community

- Papers
  - ACL Anthology has nearly everything, free!
    - Over 80,000 papers!
    - Free-text searchable
      - Great way to learn about current research on a topic
      - New search interfaces currently available in beta
        - » Find recent or highly cited work; follow citations
    - Used as a dataset by various projects
      - Analyzing the text of the papers (e.g., parsing it)
      - Extracting a graph of papers, authors, and institutions (Who wrote what? Who works where? What cites what?)
  - Google Scholar to sort by citation count / track citations
  - arXiv papers

The NLP Research Community

- Institutions
  - Universities:
  - Several “big players” with many faculty
  - Companies:
    - Old days: AT&T Bell Labs, IBM
    - Now: Microsoft Research, Google Brain/DeepMind, FAIR, Amazon, startups …
      - Nonprofits: AI2, HuggingFace, TTIC, …
      - Many niche markets – online reviews, medical transcription, news summarization, legal search and discovery …

The NLP Research Community

- Software
  - Lots of people distribute code for these tasks
    - Search github – fun to download packages and play around!
    - Or you can email a paper's authors to ask for their code
    - PapersWithCode.com
    - Search for "awesome NLP" for some lists
  - Toolkits and end-to-end pipelines for text analysis
    - Hugging Face – > 1,143,998 models, > 248,002 datasets
      - Large pretrained models: `pip install transformers` (quick tour)
      - Task-specific models: `pip install allennlp`, etc.
    - Allen NLP (Python), Spacy (Cython), UDPipe (C++), Stanza (Python), CoreNLP (Java), NLTK (Python)

Dr. Chetana is an Professor and Program lead in the CSIS department at Work Integrated Learning Programmes Division, BITS Pilani.

She has more than 27 years of teaching and industry experience. She did her PhD in Computer Science and Engineering from a joint programme of IIT Bombay and Monash University, Australia. She has been working extensively on different state of art research projects and has been awarded the “Best Industry Aligned Research” at the CSI TechNext India 2019 - Awards to Academia. She has published various papers and is also a reviewer at national and international level peer reviewed conferences and journals. Her areas of expertise include Machine Learning, Natural Language Processing, Semantic Web, Deep Learning, Text Mining, Big Data Analytics, Information Retrieval and Software Engineering.

Thank you!!

Good References

https://web.stanford.edu/~jurafsky/NLPCourseraSlides.html
https://web.stanford.edu/~jurafsky/slp3/
https://www.geeksforgeeks.org/top-7-applications-of-natural-language-processing/
https://emerj.com/partner-content/nlp-current-applications-and-future-possibilities/
https://www.nltk.org/book/
https://www.coursera.org/learn/python-text-mining/home/week/1
https://openai.com/api/
https://analyticssteps.com/blogs/top-nlp-tools
https://web.stanford.edu/~jurafsky/NLPCourseraSlides.html
https://www.cstr.ed.ac.uk/emasters/course/natural_lang.html
https://web.stanford.edu/class/cs224u/2016/materials/cs224u-2016-intro.pdf
https://www.mygreatlearning.com/blog/trending-natural-language-processing-applications/
https://grammarchecker.io/page/plagiarism-checker
https://artificialanalysis.ai/leaderboards/models

References

- Semantic Web stack of standards, W3C®
- Linking Open Data cloud diagram by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net
- https://www.geeksforgeeks.org/top-7-applications-of-natural-language-processing/
- https://levity.ai/blog/11-nlp-real-life-examples
- https://www.youtube.com/watch?v=otH29Uoo-HE
- https://www.youtube.com/watch?v=_UVfwBqcnbM
- https://www.youtube.com/watch?v=BELFUhssjXE
- Overview of NLP Tasks and Featurization Frank Ferraro – ferraro@umbc.edu
- Generative AI for everyone- coursera
- Generative AI with LLM- courser
- https://www.assemblyai.com/blog/the-top-free-speech-to-text-apis-and-open-source-engines/
- https://www.naturalreaders.com/online/
- https://www.techtarget.com/whatis/feature/12-of-the-best-large-language-models
- https://falconllm.tii.ae/