ELEC4010I / COMP4901I
Building Interactive Intelligent Systems
Final Project Overview

26.Mar.2018
http://course.ece.ust.hk/elec4010i/

Introduction:

- The goal of the final project is to let you learn how to
  - **Survey** related works in an on-going NLP research topic
  - **Analyze** real research-oriented dataset
  - **Implement** effective machine learning models
  - **Summarize** as an international conference research paper
- **2 people** per group (if you want to do project alone, please inform us first)
- **25%** in total grading
- Please download dataset and other useful files here:
  - https://github.com/hkusthltc/Final-Project
  - Please **do not** push your code online before deadline

Time Table:

- Announcement:
  - March 28 (Wed)
- Midterm report:
  - **April 18 (Wed) at 23:59**
  - **April 18 (Wed) at 23:59**
  - **April 18 (Wed) at 23:59**
- Final report:
  - **May 16 (Wed) at 23:59**
  - **May 16 (Wed) at 23:59**
  - **May 16 (Wed) at 23:59**

Submission: Midterm Report

- Deadline: April 18 (Wed) 23:59
- Send to hkust.hltc.courses@gmail.com and pascale@ece.ust.hk with the title "Midterm Report". The email should only include a **PDF** file.
- A simple report (**1 page**) including the following information:
  - Group information: Names, Student IDs, Emails
  - Topic chosen
  - Related papers (at least 3 papers)
  - Possible methodology to solve the task
  - ...

Submission: Final Report

- Deadline: May 16 (Wed) 23:59
- Send to hkust.hltc.courses@gmail.com and pascale@ece.ust.hk with the title "Final Project Report". The email should include a **PDF** file and **CODE**.
- Official ACL paper (**4-page** short paper) in **PDF** format.
  - **Abstract**
  - Introduction
  - **Methodology**
  - **Experiment Results**
  - Related Works
  - **Discussion and Conclusion**
  - Reference (unlimited pages)
- Your source code should be
  - Runable / Reproducibility

Submission: Final Report

Write your names here!

Instructions for ACL-2018 Proceedings

Anonymous ACL submission

Abstract

This document contains the instructions for preparing a camera-ready manuscript for the proceedings of ACL-2018. The document itself conforms to its own specifications, and is therefore an example of what your manuscript should look like. These instructions should be used for both papers submitted for review and for final versions of accepted papers. Authors are asked to conform to all the directions reported in this document.

1 Credits

This document has been adapted from the instructions for earlier ACL and NAACL proceedings, including those for ACL 2017 by Dan Gildea and Min-Yen Kan, NAACL-2016 by Margaret Mitchell, ACL-2012 by Maggie Li and Michael White, those from ACL-2010 by Jing-Shing Chang and Philipp Koehn, those for ACL-2008 by Johanna D. Moore, Simone Teufel, James Allan, and Sadaoki Furui, those for ACL-2005 by Hwee Tou Ng and Kemal Oflazer, those for ACL-2002 by Eugene Charniak and Dekang Lin, and earlier ACL and EACL formats. Those versions were written by several people, including John Chen, Henry S. Thompson and Donald Walker. Additional elements were taken from the formatting instructions of the *International Joint Conference on Artificial Intelligence* and the *Conference on Computer Vision and Pattern Recognition*.

2 Introduction

The following instructions are directed to authors of papers submitted to ACL-2018 or accepted for publication in its proceedings. All authors are required to adhere to these specifications. Authors are required to provide a Portable Document Format (PDF) version of their papers. **The proceedings are designed for printing on A4 paper.**

All MSWord formatting for ACL-2018 is made available in the MSWord Styles in this template. In newer versions of MSWord, click Home, then expand the Styles tile by clicking the diagonal arrow on the lower left corner. This should open all styles in the template for you to apply to your document as needed. Otherwise, you may expose the Styles following the instructions provided at: http://blogs.technet.com/b/hub/achive/2010/11/22/view-and-edit-styles-quickly-in-word-2010.aspx.

3 General Instructions

Manuscripts must be in two-column format. Exceptions to the two-column format include the title, authors' names and complete addresses, which must be centered at the top of the first page, and any full-width figures or tables (see the guidelines in Subsection 3.6). Lines should be justified, with even spacing between margins (Ctrl+J). Single-spaced lines are permitted, but **authors are encouraged to use Paragraph spacing at Multiple, 1.05pt**, with **Font character spacing condensed with kerning of 0.1pt**, and **Margins at 0.98in**, for consistency with A4 paper and documents formatted with LaTeX. Go to Format, Document, Page Setup, and make sure A4 is selected. The manuscript should be printed single-sided and its length should not exceed the maximum page limit described in Section 5.

Pages have a ruler, header and footer for initial submission, with header at 0.3 in from top and footer at 0.4 in from bottom. When you first create your submission on softconf, please fill in your submitted paper ID where *** appears in the *** at the header at the top. If the paper is accepted, remove the header, footer (page numbers), and the ruler for the final version (camera-ready). **Do not number the pages in the camera-ready version.**

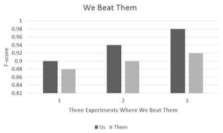


Figure 1: Figure caption.

Placement: Place figures and tables in the paper near where they are first discussed, as close as possible to the top of their respective column. Wide figures and tables may run across both columns and should be placed at the top of a page.

In MSWord, authors can place a Figure (*e.g.*, a graphic and its caption) inside the rows of a 2 x 1 table (2 rows and 1 column) with invisible borders. Specify table positioning by right-clicking its handle in the upper left corner. Place the image in the center of the first row, and the caption in the center of the second row.

Captions: Provide a caption for every table and figure; number each one sequentially in the form: "Figure 1: Figure caption.", "Table 1: Table caption." Type the captions of the figures and tables below the body, using 11-point font.

Numbering: To update numbering, highlight all the relevant text (*e.g.*, **Ctrl-A + F9**). This will update all the numbering applicable to tables, figures, equations, and headings.

Cross-referencing: To add a cross reference to a figure or table:

- Place the mouse pointer at the location where you wish to add the cross-reference.
- Click on the **Insert** menu, (then click **Reference**), and then **Cross-reference** in the **Links** panel.
- In the **Cross-reference** dialog box, click the caption to which you are building the text reference.
- For a figure, under **Reference Type**, click **Figure**.
- Under Insert Reference To, click Only Label and Number, then click OK.
- Once the reference is in place, apply the 'Normal' font style (size 11, no bold face).
- This is an example reference to Figure 1.

3.10 Equations

An example equation is shown below:

$$A = \pi r^2 \quad (1)$$

To add new equations, authors are encouraged to copy this existing equation line, and then replace with the new equation. The numbering and alignment of equation line elements is automatic. To update equation numbering, press **Ctrl-A + F9**. Note: this will only update the number to the right of the equation; to update numbering within the text you must create a cross-reference.

Cross-referencing: To create a cross-reference for an equation:

- Create a bookmark for it.
- Select the number to the right of the equation. Go to **Insert, Bookmark** (in the **Links** panel), and then create a name for your equation. Press **Add** to create the bookmark.
- To refer back, place the mouse pointer at the location where you wish to add the cross reference.
- Go to **Insert, Cross-reference** (in the **Links** panel). In the dialogue box, select **Bookmark** and **Bookmark Text** from each dropdown list. Uncheck **Insert as Hyperlink**, then click **OK**.
- This will make it such that whenever a new equation is added, the references to the equation will be updated when **Ctrl-A + F9** is pressed.
- This an example cross-reference to Equation (1).

3.11 Accessibility

In an effort to accommodate the color-blind (as well as those printing to paper), grayscale readability for all accepted papers is encouraged. Color is not forbidden, but authors should ensure that tables and figures do not rely solely on color to convey critical distinctions. A simple criterion: All curves and points in your figures should be clearly distinguishable without color.

3.12 From Submitted to Camera-Ready

To remove submission formatting for the camera-ready document, delete the header and footer on

Submission: Final Report

- Deadline: May 16 (Wed) 23:59
- Send to hkust.hltc.courses@gmail.com and pascale@ece.ust.hk with the title "Final Project Report". The email should include a **PDF** file and **CODE**.

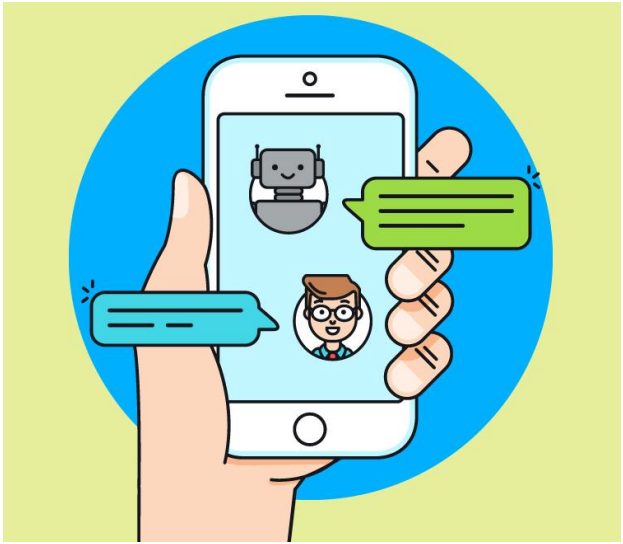
- What are the difference between final project and homeworks?
  - **Data**: You have all the labels for train, validation and test sets.
  - **Baselines**: There are other baselines available online that you can compare with. You can also make your own baselines.
  - **Methodology**: In homeworks you have the instructions step-by-step, but here you need to figure out what to do yourself.
  - **Evaluation**: Although we will provide one way for each task to evaluate your results, you can also find your own way to show your strength.

Three different Topics (Choose 1)

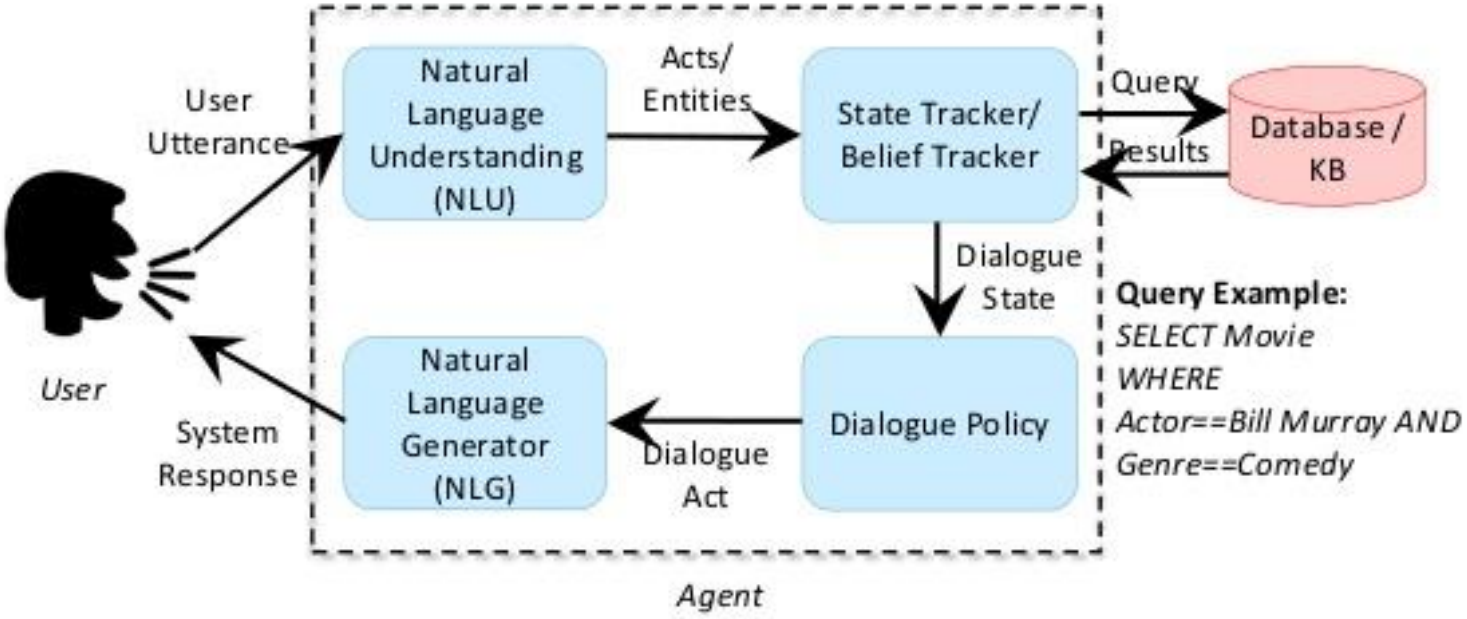
- Task 1: Goal-Oriented Dialog Learning
- Task 2: Emotion Analysis
- Task 3: Machine Translation

Three different Topics

- Task 1: Goal-Oriented Dialog Learning
- Task 2: Emotion Analysis
- Task 3: Machine Translation



Task 1: Goal-Oriented Dialog Learning



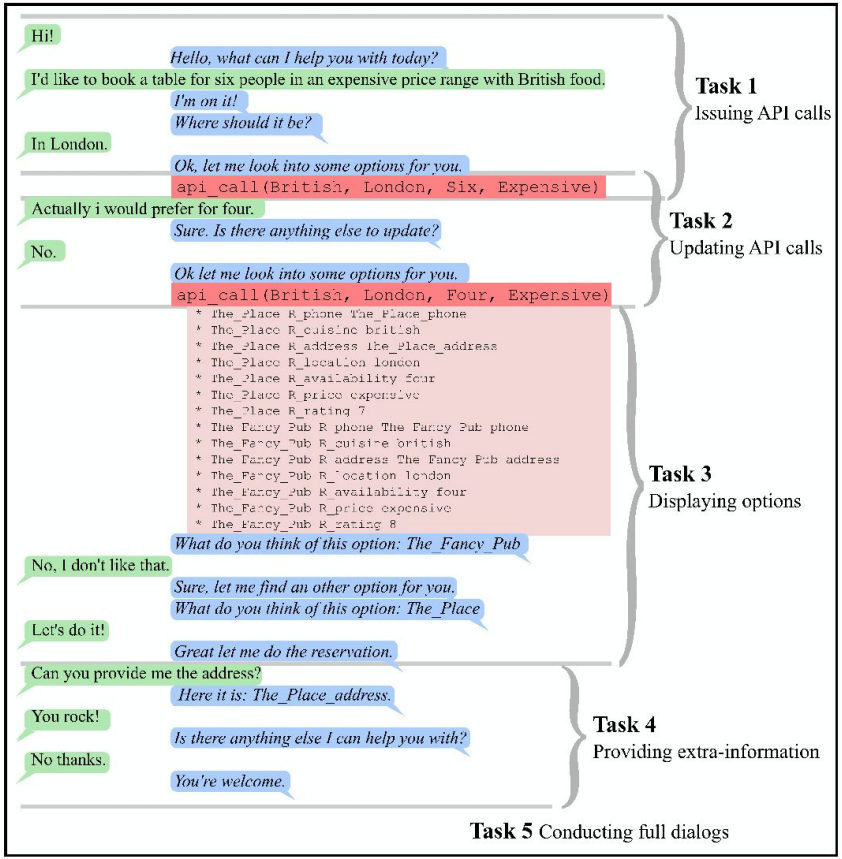
Task 1: Goal-Oriented Dialog Learning

Task 1: Goal-Oriented Dialog Learning

- Goal-oriented dialog technology is an important research issue and End-to-End dialog learning has emerged as a primary research subject in the domain of conversation agent learning.
- It consists in learning a dialog policy from transactional dialogs of a restaurant reservation domain.
- This task is from DSTC-6 (Dialog System Technology Challenges) series, which aims to build End-to-End dialog systems for goal-oriented applications.
- Officially there are 4 test sets, here you only need to solve **test set 1**.

Task 1: Goal-Oriented Dialog Learning

- Five tasks:
  - Task 1: Issuing API calls
  - Task 2: Updating API calls
  - Task 3: Displaying options
  - Task 4: Providing extra information
  - Task 5: Conducting full dialogs



Task 1: Goal-Oriented Dialog Learning

- Five tasks, 10000 dialogs each task, 10 candidates each dialog
- **Rank** your candidates for each dialog

```
[
        {dialog_id : " ",
        utterances: [" ", ... ],
        candidates: [
                {candidate_id: " " , utterance: ""}, ...
                ]
        }
]
```

```
[
        {dialog_id : " ",
        lst_candidate_id:
                [ {candidate_id: " ", rank: " "}, ... ]
        }
]
```

Task 1: Goal-Oriented Dialog Learning

dataset_walker.py Parse the JSON file and print out a plain version of the dataset comprising the dialog utterances, the candidate answers and the correct answers for each dialog in the corpus.

check_validity.py Load a result file and perform a series of verification regarding the format of the result file.

score.py Compute the score of a candidate result file compared to a reference. This is the script that will be used to evaluate the participants during the testing phase of the challenge.

baseline_random.py Produce a random baseline result file for a given task dataset file.

baseline_tfidf.py Produce a baseline result file for a given task dataset file using a TF-IDF similarity heuristic between the candidate answers and the corresponding dialog.

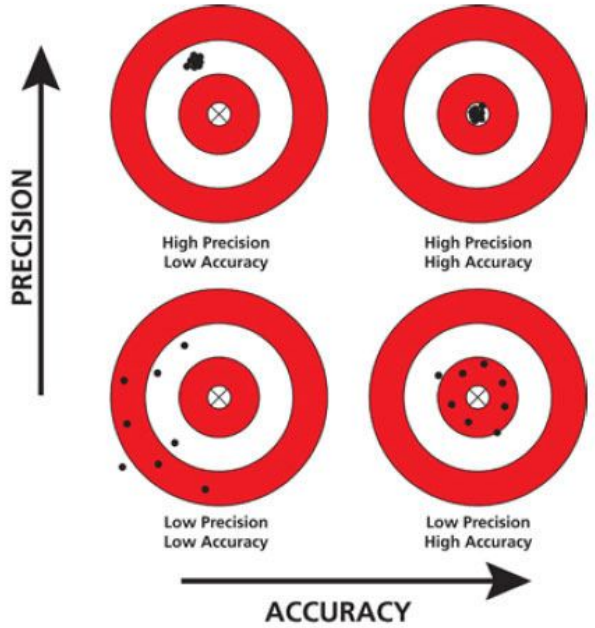
Task 1: Goal-Oriented Dialog Learning

```
dialog_id: d2f95f9f-ffb1-446e-9337-334f41ecc210
Utterances:
 * hi
 * hello what can i help you with today
 * may i have a table for four in paris with a romantic atmosphere in a moderate price range
 * i'm on it
 * <silence>
 * any preference on a type of cuisine
 * sorry one sec, i am asking my friend and she wants british, let's do that
 * ok let me look into some options for you
 * <silence>
Candidates:
 * d363cb37-9398-475a-985d-35d0c54dc75c - api_call french london eight expensive casual
 * 948ae1aa-60ab-4cc6-a833-3f7c8a825616 - api_call spanish paris two moderate romantic
 * 9bfbb7d5-caf0-43c0-b9c9-1119cc35e60d - api_call british paris four moderate romantic
 * ea4440ef-cf19-4546-b02c-51bf1e6db691 - api_call british bombay four moderate romantic
 * 39df25ad-9d4c-4c2f-9079-e4407965ad48 - api_call italian paris four moderate romantic
 * 4c3fe95a-2ba3-40b3-b04e-0f731735eea1 - api_call british paris two moderate romantic
 * 886196db-e824-45d3-af92-d9ab0626ff2c - api_call french paris eight moderate romantic
 * 1a038abc-3c01-4104-b5b9-c66d175dad04 - whenever you're ready
 * a3a36272-3b11-48f9-8544-8249f5df50c3 - api_call british rome four moderate casual
 * b7e16ac7-19d0-4002-bdf1-bd1ccf707bc1 - api_call british london four moderate romantic
Answer:
 * 9bfbb7d5-caf0-43c0-b9c9-1119cc35e60d - api_call british paris four moderate romantic
```

Task 1: Goal-Oriented Dialog Learning

- Evaluation
  - Precision@1
  - Precision@2
  - Precision@5

- Check score.py for more informaiton

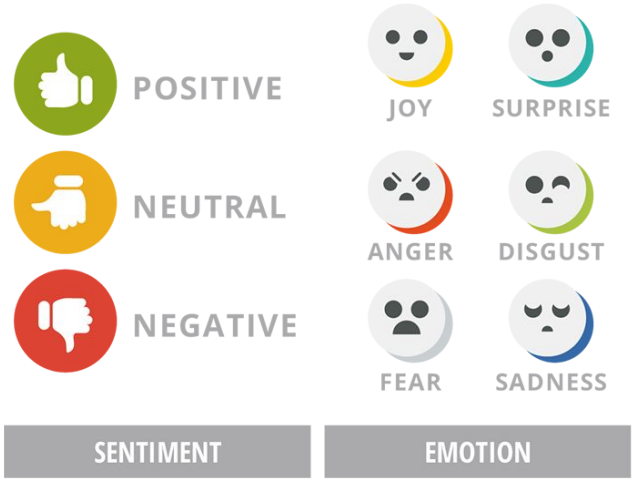


Task 1: Goal-Oriented Dialog Learning

- Recommended Reading
  - Generative Encoder-Decoder Models for Task-Oriented Spoken Dialog Systems with Chatting Capability
    - https://arxiv.org/abs/1706.08476
  - Hybrid Code Networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning.
    - https://arxiv.org/abs/1702.03274
  - Learning end-to-end goal-oriented dialog.
    - https://arxiv.org/abs/1605.07683

Three different Topics

- Task 1: Goal-Oriented Dialog Learning
- Task 2: Emotion Analysis
- Task 3: Machine Translation



Task 2: Emotion Analysis

- SemEval-2018 Affect in Tweets DIstant Supervision Corpus (SemEval-2018 AIT DISC)
- This corpus of tweets was collected by polling the Twitter API for tweets that included emotion-related words such as 'angry', 'annoyed', 'panic', 'happy', 'elated', 'surprised', etc.
- Officially there are 5 subtasks, here you only need to solve **task 1** and **task 2**.

Task 2: Emotion Analysis

- Two tasks:
  - task A: Emotion Intensity
  - task B: Emotion Classification
- Four emotion:
  - Joy, Angry, Sadness, Fear

EI-reg, EI-oc	Train	Val	Test	Total
anger	1,701	388	1,002	3,091
fear	2,252	389	986	3,627
joy	1,616	290	1,105	3,011
sadness	1,533	397	975	2,905

Task 2: Emotion Analysis

ID	Tweet	Affect Dimension	Intensity Score
2018-En-02328	@PageShhh1 I know you mean well but I'm offended. Prick.	anger	0.734
2018-En-02617	Let go of resentment, it will hold you back, do not worry about what could of been, what is to come is what matters.	anger	0.422
2018-En-01021	No, I'm not 'depressed because of the weather,' I'm depressed because I have #depression #sicknotweak	anger	0.663
2018-En-03737	#AmarnathTerrorAttack Muslims are killing everywhere Syria Iraq Palestine Everyday beyond They say that Islam is terrorism shame on you	anger	0.703
2018-En-03407	Prepare to suffer the sting of Ghost Rider's power! Prepare to know the true meaning of hell!	anger	0.719
2018-En-02897	@ajduey04303 We've been broken up a while, both moved on, she's got a kid, I don't hold any animosity towards her anymore...	anger	0.359
2018-En-04119	Just know USA, all Canadians don't agree with what Khadr's settlement and his unwillingness to take responsibility for his actions. #outrage	anger	0.844

ID	Tweet	Affect Dimension	Intensity Score
2018-En-00290	People are truly #amazing. #inspiring day	joy	0.712
2018-En-03093	What are some good #funny #entertaining #interesting accounts I should follow ? My twitter is dry	joy	0.339
2018-En-00588	@NoahWebHouse A review of my book faulted me for spending so much time on Webster's introductory front matter. But I had to, #brilliant!	joy	0.533
2018-En-01532	#good to learning #wisdom << reform (v): make in order to improve something >>	joy	0.281
2018-En-03258	Really excited to see our team this year, and especially moving in the next 2-3 years when his guys start playing significant minutes.	joy	0.797
2018-En-03332	Be happy. Be confident. Be kind.\n\n #KissablesLoveSMShopmag\nAllOutDenimFor KISSMARC	joy	0.661

ID	Tweet	Affect Dimension	Intensity Score
2018-En-02626	@EdmundPAdamus That's the horrific irony; while we 'cleanse' society of moral value, a new religion will rise up & fill the gap - unhindered	fear	0.596
2018-En-00320	@TuckerCarlson @krauthammer @FoxNews They always need the drama and a fight to feel adaquate. Sad power is so insecure!!!	fear	0.533
2018-En-03973	@executivegoth I tried a free Alpha trial to watch #dread and I'm sold. Worth the subscription alone. You all are incredible!	fear	0.179
2018-En-02615	@anya_sisi Off to jail with lots of other like minded muslims, with his koran, prayer mat, halal #islam = #terrorism	fear	0.640
2018-En-03312	My boss said to another girl who is leaving 'i hope your new job is shit.' How pathetic. Hence why people are leaving in droves.	fear	0.533
2018-En-02333	@cnn please get Sebstian Gorka off of my tv. The whole trump admin. needs to stop deflecting by talking about Hillary.	fear	0.328
2018-En-00741	ever shit and just become another cog in the chat i will not hesitate to smack ur bottoms	fear	0.444

ID	Tweet	Affect Dimension	Intensity Score
2018-En-03239	@TerraJole is a bully. plain and simple.	sadness	0.550
2018-En-02839	@LondonEconomic Sometimes our judiciary just leaves you breathless and speechless.	sadness	0.521
2018-En-01035	#pause You gotta luv March Madness. Mad?? More like #upset	sadness	0.525
2018-En-01920	Massive night tonight with the decider! Who you got? @QLDmaroons or @NSWRL ??? Can the #blues get it done?! #origin #StateOfOrigin	sadness	0.306
2018-En-03155	@RGUpdate Have you tried English hospital food ?? #yak #gross #horrible	sadness	0.553
2018-En-01336	Morning Swindon!\nIs there any chance you can cheer yourself up a bit?!! #bleak 😝	sadness	0.469

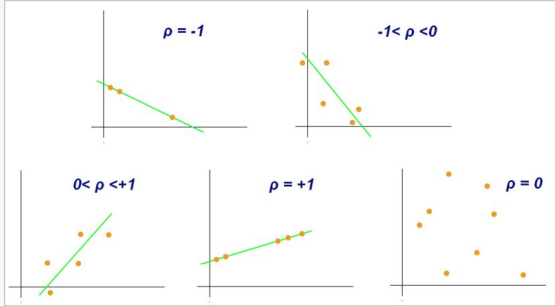
Task 2: Emotion Analysis

- Evaluation: Pearson Correlation Coefficient with the Gold ratings/labels
- https://en.wikipedia.org/wiki/Pearson_correlation_coefficient

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \qquad \rho_{X,Y} = \frac{\operatorname{E}[XY] - \operatorname{E}[X]\operatorname{E}[Y]}{\sqrt{\operatorname{E}[X^2] - [\operatorname{E}[X]]^2}\ \sqrt{\operatorname{E}[Y^2] - [\operatorname{E}[Y]]^2}}$$

where:

- cov is the covariance
- σ_X is the standard deviation of X
- σ_Y is the standard deviation of Y



Examples of scatter diagrams with different values of correlation coefficient (ρ)

Task 2: Emotion Analysis

- Recommended Reading
  - Convolutional Neural Networks for Sentence Classification
    - https://arxiv.org/abs/1408.5882
  - Emotion Recognition on Twitter: Comparative Study and Training a Unison Model
    - http://ieeexplore.ieee.org/document/8295234/

Three different Topics

- Task 1: Goal-Oriented Dialog Learning
- Task 2: Emotion Analysis
- Task 3: Machine Translation

Task 3: Machine Translation

- Statistical machine translation (SMT)
  - Distributional Semantics
    - Quantifying and categorizing semantic similarities between linguistic items based on their distributional properties in large samples..
    - Distributional hypothesis: linguistic items with similar distributions have similar meanings
- Neural machine translation (NMT)
  - Use a artificial neural network to predict the likelihood of a sequence of words
  - Its main departure is the use of vector representations ("embeddings", "continuous space representations") for words and internal states

Task 3: Machine Translation

- Data:
  - (Suggested) IWSLT'15 English-Vietnamese data
    - Around 133K training sentence pairs
    - Around 1300 testing sentecne pairs
  - Others (If you have more time and GPUs)
    - WMT'14 English-German data
      - Around 4.5M training sentence pairs
    - WMT'15 English-Czech data
      - Around 15.8M training sentence pairs

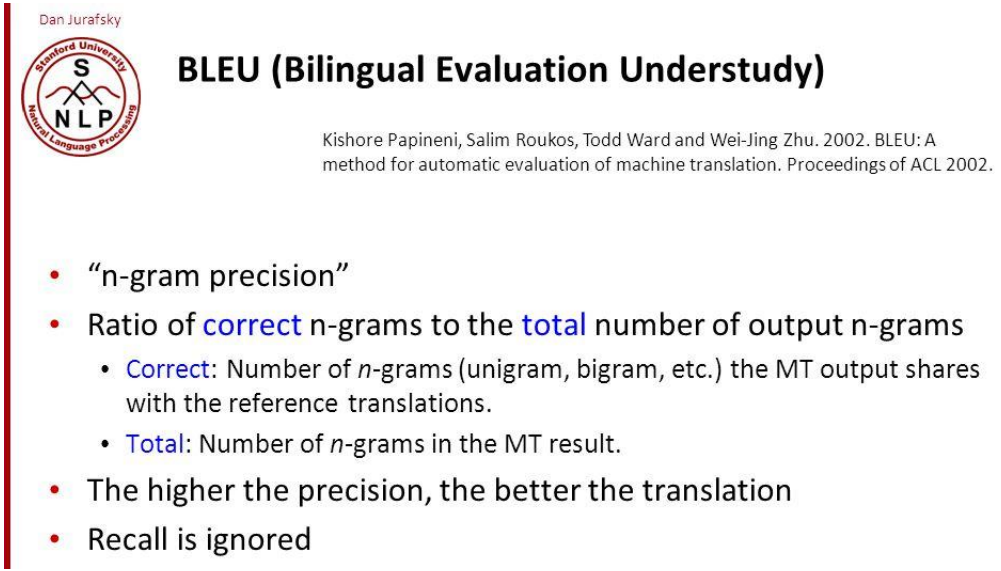
Task 3: Machine Translation

- The weight of a paper clip is approximately equal to 900 zeta-illion -- 10 to the 21st -- molecules of isoprene .
- But despite its very small weight , enough of it is emitted into the atmosphere every year to equal the weight of all the people on the planet .
- It 's a huge amount of stuff . It 's equal to the weight of methane .
- And because it 's so much stuff , it 's really important for the atmospheric system .
- Because it 's important to the atmospheric system , we go to all lengths to study this thing .
- We blow it up and look at the pieces .
- This is the EUPHORE Smog Chamber in Spain .
- Atmospheric explosions , or full combustion , takes about 15,000 times longer than what happens in your car .
- But still , we look at the pieces .

- Trọng lượng của một chiếc kẹp giấy vào khoảng 900 zeta-illion -- 10 mũ 21 -- phân tử isoprene .
- Dù trọng lượng phân tử rất nhỏ , thế nhưng lượng isoprene được thải vào khí quyển hàng năm ngang ngửa với tổng trọng lượng của dân số toàn cầu .
- Đó là một lượng khí thải khổng lồ , bằng tổng trọng lượng của mêtan .
- Chính vì lượng khí thải rất lớn , nó có ý nghĩa quan trọng với hệ thống khí quyển .
- Chính vì nó có ý nghĩa quan trọng với hệ thống khí quyển , giá nào chúng tôi cũng theo đuổi nghiên cứu này đến cùng .
- Chúng tôi cho nó nổ và xem xét từng mảnh nhỏ .
- Đây là Phòng nghiên cứu khói bụi EUPHORE ở Tây Ban Nha .
- Nổ trong không khí hay cháy hoàn toàn diễn ra chậm hơn 15,000 lần so với những phản ứng trong động cơ xe .
- Dù vậy , chúng tôi vẫn xem xét từng mảnh nhỏ .

Task 3: Machine Translation

- Evaluation:
  - **B**i**L**ingual **E**valuation **U**nderstudy (BLEU) score
  - https://en.wikipedia.org/wiki/BLEU



Task 3: Machine Translation

- Recommended Reading
  - Stanford Neural Machine Translation Systems for Spoken Language Domains
    - https://nlp.stanford.edu/pubs/luong-manning-iwslt15.pdf
  - Achieving Open Vocabulary Neural Machine Translation with Hybrid Word-Character Models
    - https://nlp.stanford.edu/pubs/luong2016acl_hybrid.pdf
  - Neural Machine Translation by Jointly Learning to Align and Translate
    - https://arxiv.org/abs/1409.0473

Other Tasks:

- Students can propose other topics they are interested in and related to the course. It should be an on-going research topics and should get an approval from professor first.
- Send a proposal to hkust.hltc.courses@gmail.com and pascale@ece.ust.hk with the title "Final Project Proposal".
  - Deadline: **April 11** (Wed) 23:59
  - One week before midterm report

Evaluation:

- Professor and TAs will seprately review your report based on ACL reviewer guideline
  - https://www.aclweb.org/adminwiki/index.php?title=ACL_Reviewer_Guidelines
- A sample review form will be like this:
  - http://acl2018.org/downloads/acl_2018_review_form.html

Evaluation:

- Strengths and Weakness
- Originality
- **Soundness/Correctness**
- Substance
- **Replicability**
- Meaningful Comparison
- **Readability**

General Hints

- **Understand** your problem
- **Preprocess** your data
  - data cleaning
  - tokenization, build a word2idx dictionary, vectorize, etc
  - data statistics
- **Design** and **evaluate** your model
  - Find a model that is suitable for your task and implement it. At the same time, you should have some baselines
  - Tune some parameters / find the best one
  - Find the weakness, rethink about the problem

- Please feel free to come during office hours to **discuss with TAs**…

Rules:

- All the ML libraries online are available to use
  - PyTorch, Sklearn, …
- **No Plagiarism**
  - No credit if we find out cheating
- No late submission except special cases informed in advance
  - **10%** deduction for each day

Enjoy!