# BERT-Supervised Encoder-Decoder for Restaurant Summarization with Synthetic Parallel Corpus

**Lily Cheng**
Stanford University CS224N
lilcheng@stanford.edu

## Abstract

With recent advances in seq-2-seq deep learning techniques, there has been notable progress in abstractive text summarization. However, the difficulty in obtaining large corpus of document-summary pairs remains a barrier to the application of supervised learning techniques in new domains. A new task of generating restaurant summaries from online articles is proposed under a low resource regime. Concretely, *no human-generated summaries* are used in training. A novel extractive-generative model to train a text summarizer using synthetic seq-2-seq pairs is proposed. We leverage on successes in transfer learning to supervise the extraction stage, which then guides the text generation. Results so far demonstrate promise at the sentence level, proving that it is possible to generate sensible output for a new task under resource constraints. Analysis of common errors point to the need for further architectural modifications and tests to reach commercial-grade results.

## 1 Introduction

### 1.1 The Low-Resource Challenge

In recent years, abstractive multi-document summarization has progressed rapidly, drawing from advances in neural machine translations. The availability of the CNN/Daily Mail dataset[8] with 286k document-summary pairs has propelled research beyond the realm of sentence compression or single-sentence summaries from short paragraphs.

However, abstractive summarization remains challenging to apply to new tasks due to the large in-domain document-summary corpus required. Successful application of transfer learning to out-of-domain tasks has been limited. This is a significant impediment to mainstream adoption of abstractive summarization techniques beyond the realm of academia and big tech firms.

This project aims to answer: "Is it possible to conduct abstractive summarization on new domains without using human-generated summaries for training?"

### 1.2 Proposed New Task With Low Resource: Restaurant Summarization

To answer this question, a real-world task of generating restaurant summaries from online articles such as the Michelin Guide is proposed. No existing in-domain document-summary corpus is available for supervised learning. Through novel use of transfer learning together with synthetic corpus generation, we evaluate the feasibility of this strategy to tackle low-resource tasks. If viable, this strategy can be adapted to other domains.

A important unique point to this project is that it was conducted with only annotating 1500 sentences from the Michelin Guide using only 4 hours of human time and, unlike established abstractive summarization techniques, no human-generated summaries were produced for training.

## 1.3 Key Idea and Novel Contributions

The project introduces a two-stage extractive-generative model. The first stage leverages transfer learning with BERT to conduct a phrase and named-entity extraction. The second stage uses the output from the first stage, with the original text to produce a synthetic corpus to train the encoder-decoder stage. Instead of the typical human-supervised seq-2-seq training, we have a BERT-supervised training.

This is partially inspired by the work of Sennrich et al. [11] where synthetic parallel data was created using the concept of back-translation to improve NMT for low resource language pairs.

## 1.4 Results

Within the time available for this project, it has been demonstrated that even with a very small amount of annotated text and no human written summaries, it is feasible to train a model to generate text that can be natural and coherent at a sentence level as shown below. However, it was much more challenging to achieve any degree of consistency in the generation of a multi-sentence summary.

As an example of the abstractive properties of the model, the summary sentence below uses the word "specialties" which was not word present in the original text.

| |
|---|
| **Original Text**: In the beautiful old town house dating from 1787 there is a modern bistro on one side and a tasteful historical atmosphere on the other - an interior well work seeing! The menu at this modern yet cosy restaurant includes dishes such as rib steak with onions, mustard and horseradish and pan-fried new potatoes. |

| |
|---|
| **Summary**: This beautiful old town house has a modern bistro with a tasteful historical atmosphere. A modern yet cosy restaurant with specialities such as rib steak with onions, mustard and horseradish or pan-fried new potatoes. |

## 2 Related Works

**Abstractive Summarization:** Neural models have been used for abstractive summarization at the sentence level (Rush et al., 2015 [9], Chopra et al., [2]) using the Gigaword dataset of pairing the first sentence of each news article with the headline as the target summary. This was extended to multi-sentence summaries with the CNN/DailyMail corpus (Nallapati et al., 2016 [8]). Pointer-generator network proposed (See et al., 2017 [10]) further addressed problems dealing with Out-of-vocabulary (OOV) words and repetitiveness, combining an extractive and abstractive approach. In a similar but more explicit extractive approach (Li et al., 2017 [5], TextRank algorithm is used to extract important keywords which is encoded and passed to decoder.

**Transfer Learning in Natural Language Understanding:** BERT (Devlin et al., 2018) achieved significant progress in transfer learning for natural language understanding using the transformer architecture (Vaswani et al., 2017 [12]). A pre-trained BERT fined-tuned with an additional output layer was proven to deliver state-of-the-art results for a large range of natural language understanding tasks. The extractive phase of our model is based on BERT.

**Encoder-Decoder Neural Models:** Attention-based encoder-decoder networks (Luong et al., 2015 [7]) have been successfully used for neural machine translation and can be used more generally for seq-2-seq tasks. To address OOV challenges, a word-character model is proposed (Luong et al., 2016 [6]) replacing the <unk> symbol with a model that works at the character-level. In the E2E NLG Challenge (Dusek et al., 2018 [3]), the leading model (Juraska et al., 2018 [4]) used a similar attention-based encoder-decoder RNN model for natural language generation based on (Bahnadau et al., 2014 [1]). Delexicalization was used to improve the generalizability of the model and reduce amount of training data needed.

**Synthetic training:** To address problems in NMT for low resource language pairs, synthetic parallel data was created using back-translation (Sennrich et al., 2015 [11]). We draw on the concept of synthetic parallel data to train the encoder-decoder stage of our model.

# 3 Overall Approach and Model Architecture

## 3.1 Pipeline Overview

This 4-step process is most easily understood with reference to diagram below.
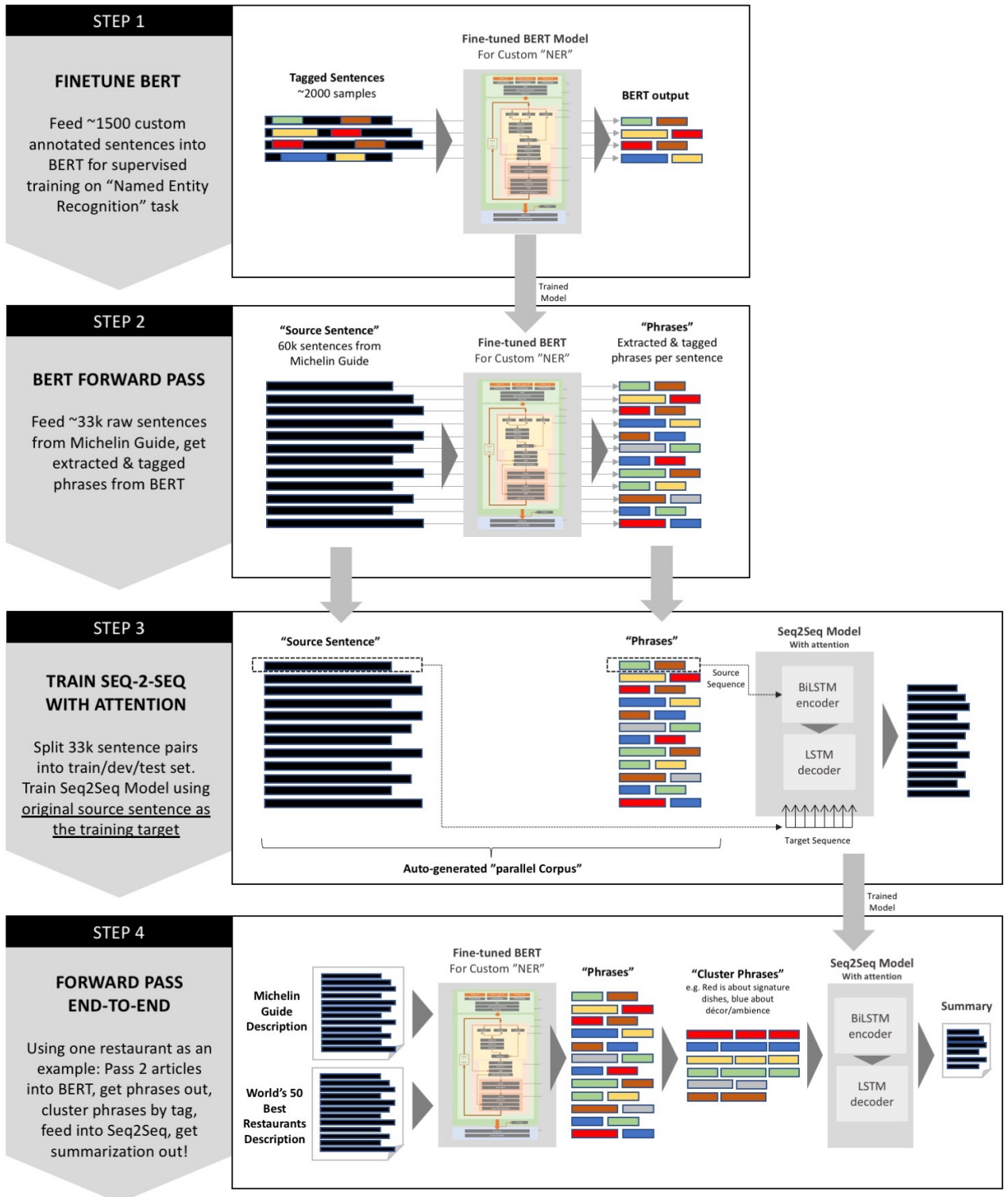
## 3.2 Phrase Extraction Model with BERT

For extracting important named-entities and phrases from the source text, pre-trained BERT[1] is used with an additional linear output layer with 11 output features, corresponding to the number of classes. The Cross Entropy Loss is used between the gold labels and the logit scores from BERT. Please refer to Appendix A for an architectural diagram of BERT and the additional layer added. This is fined-tuned using 1500 custom annotated sentences from the Michelin Guide.

$$loss(x, gold\_label) = -log(\frac{exp(x[gold\_label])}{\sum exp(x[j])})$$

## 3.3 Encoder-Decoder Model for Abstractive Generation

For our encoder architecture, we use a bi-directional single-layer LSTM. For OOV words, we use a character-based LSTM decoder in our base line model.[2] In our experiments, we also test a modified pointer feature - as illustrated below.

## 3.4 Simple Pointer-Generator mechanism

In our testing, we found that in low resource regime, the character decoder was not sufficient to accurately reproduce rare words - such as a chef's name. As such, we also implemented a simple pointer mechanism to switch between copying the source word or using the character-based generator.

$$P_t^{char\_gen} = \Big[ \prod_{i=0}^{L} P_{char}(i) \Big]^{1/L}$$

$$\alpha_t = Softmax\big[(h_t^{dec})^T W_{attProj} h_i^{enc}\big]$$

$$P_t^{att} = \frac{Max(\alpha_t)}{\sum_{i=0}^{S} \alpha_t}$$

For a given timestep $t$, $P_t^{char\_gen}$ is the probability of character decoder word. This is calculated by taking the product of the probabilities of each individual character normalized for length of word. $\alpha_t$ is attention distribution over the source sentence. $P_t^{att}$ is the probability mass on the source word which has the highest attention on this particular timestep. The pointer will copy from highest attention word from sentence if $P_t^{att} > P_t^{char\_gen}$ and if the highest attention source word is not a symbol. When the attention is on the comma between extracted phrases, this is when the model is being creative generating interconnecting words - we do not want it to interfere at these times.
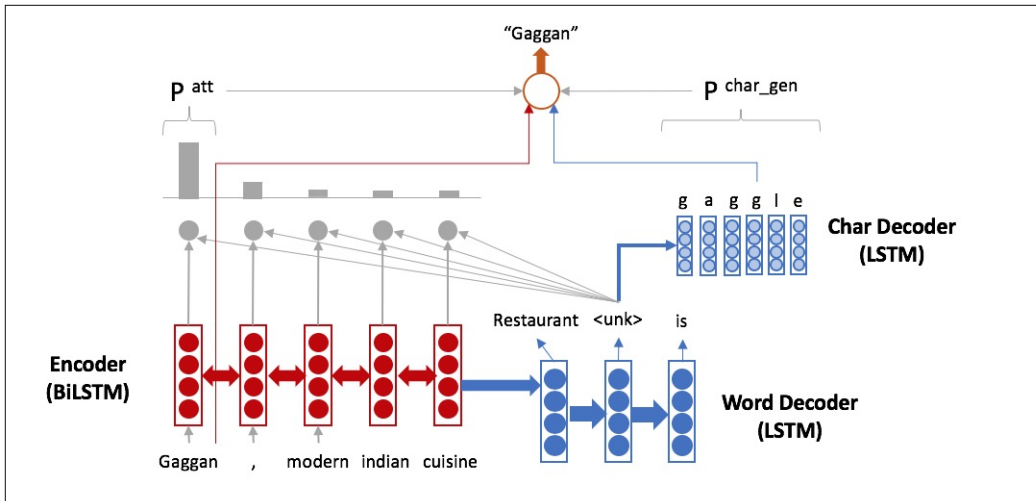


---

# 4 Experiments and Results

## 4.1 Data Preparation in STEP 1

The source text came from crawling the Michelin Guide website for restaurant descriptions. The data was cleansed by filtering out non-english text and removing any duplicate sentences. Although the source text was in english, there were a significant of number of non-standard characters (for example, french chef names). Hence the character-based CNN model was modified and trained with a more comprehensive set of characters.

The sentences were tokenized into words, automatically tagged with POS tags using the Spacy library and 1500 sentences were manually annotated with 11 custom labels as follows. For more visual examples, please refer to Appendix B:

Table 1: Custom NER labels for phrase/word extraction

| Tag | Description | Examples of tagged phrases/words |
|-----|-------------|----------------------------------|
| NAME | Name of Restaurant | *Jihwaja | Ichiryu | Thai Orchid, The Aqua* |
| TYPE | Type of food establishment | *chic brasserie | sophisticated Japanese restaurant* |
| CUISINE | Type of food offered | *traditional French cuisine | seafood restaurant* |
| DECOR_DESC | Decor/ambience description | *lifestyle space | lovely atmosphere | elegant | cosy* |
| FOOD_DESC | Food description | *classic | bold | sophisticated | daring* |
| DISH_NOUN | Signature dishes | *slow cooked beef with polenta | foie gras terrine* |
| DRINKS_NOUN | Signature drinks | *cocktails | sake | umeshu plum wine | real ales* |
| CHEF | Chef's name | *chef , Björn Freitag | chef Tom Aikens* |
| LOCATION | Location description | *opposite the British Museum | Borough Market* |
| HISTORY | Establishment history | *20 years | 1921 | since 1981* |
| SERVICE | Service description | *attentive | meticulous | warm* |

## 4.2 Step 1: BERT NER results

In this first pass, we finetuned BERT on phrase extraction task with set of 1350 tagged sentences for training and 150 sentences for evaluation. The test set had 5312 tokens of which 1146 tokens were tagged with one of the 11 custom tags.

| | Gold | Recall | | Precision | |
|---|---|---|---|---|---|
| Tag | Tokens | Token | Type | Token | Type |
| NAME | 55 | 100% | 71% | 66% | 66% |
| TYPE | 45 | 100% | 93% | 88% | 84% |
| CUISINE | 97 | 92% | 90% | 74% | 58% |
| DECOR_DESC | 250 | 90% | 85% | 58% | 56% |
| FOOD_DESC | 91 | 89% | 47% | 39% | 37% |
| DISH_NOUN | 386 | 98% | 98% | 83% | 76% |
| DRINKS_NOUN | 14 | 86% | 86% | 75% | 75% |
| CHEF | 94 | 100% | 89% | 77% | 77% |
| LOCATION | 98 | 94% | 90% | 49% | 44% |
| HISTORY | 14 | 100% | 100% | 60% | 47% |
| SERVICE | 2 | 100% | 100% | 100% | 100% |
| ALL TAGS | 1146 | 95% | 87% | 68% | 62% |

Table 2: BERT NER Task - Phrase Extraction Metrics

Despite the small training set, the recall is excellent. The "Token" recall rate refers to only the positions of the labels. The "Type" recall refers to both the position and the label type. 95% of labels were correctly positioned on the right token and 87% were correctly positioned and labelled. The precision metrics were weaker, meaning that the model tended to tag tokens which weren't tagged in the reference sentence.

A detailed investigation into the low precision scores is provided in the Analysis section.

### 4.3 Step 3: Encoder-Decoder training results

A synthetic supervised dataset generated by BERT in inference mode was used to train the encoder-decoder model.

> **Source:** tranquil area of Samcheon-gong, charming courtyard, Italian-inspired cuisine
>
> **Target:** In the tranquil area of Samo-nen's charming courtyard, this restaurant serves Italian-inspired cuisine.

The total dataset was 33k sequence pairs split into train/dev/test sets in the ratio of 80/10/10. A batch size of 32 was used, along with a learning rate of 0.001, a learning rate decay of 0.5 and was trained to 30 epochs. The graph shows that by 10k iterations, the model was starting to overfit. Given the poor perplexity score, it is clear that there is potential to significantly improve the model performance by providing more data.



Perplexity of train and dev set during training
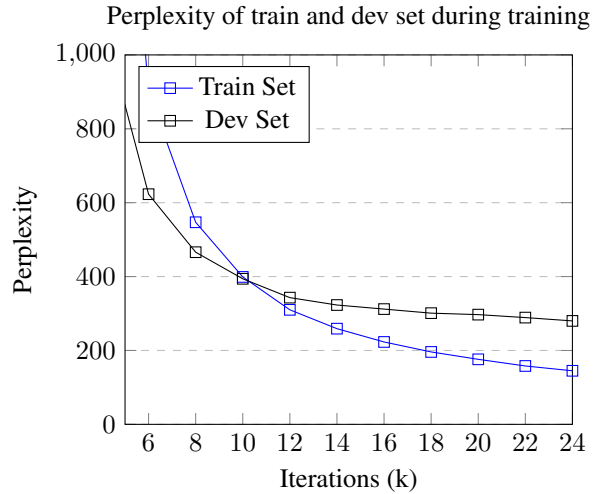
**Evaluation Metrics:** This stage was evaluated based on three key metrics. Given that synthetic nature of the data and hence there is no "gold" target per-se, the input phrases is used as our reference:

- **Phrase-level Recall:** This is defined as the percentage of input phrases incorporated into the generated sentences. Our aim is for this to be as high as possible.
- **Word-level Precision:** This is defined as the percentage of total generated words are not equal to the input phrases. Our aim is for this to be around 30-50%. This is, in essence, a measure of abstractiveness as these are all the words that are generated without "copying" from the input phrases.
- **Word-level Repeat:** we measure what percentage of words are redundant as they are repetitions of input phrases. With the introduction of the "pointer" feature, in an attempt to improve recall, the rate of repeated input phrases increased. This metric helped to measure the efficacy of the "coverage" feature in reducing repetition.
- **Human Evaluation:** 30 sentences from each set were categorized into "good", "understandable with errors", "bad by humans qualitatively. "Human Good" is good/total

**Key Experiments:** The baseline vanilla model a BiLSTM encoder and LSTM decoder with Attention combined with Character-LSTM generation for unknown words (similar to assignment 5).

- **Pointer Switch:** To improve the model's ability to copy key phrases, especially ones with rare words, a pointer switch was added to the vanilla model for testing. This increased the recall rate from 69% in the vanilla model to 74% in the pointer model.
- **Coverage:** With the introduction of the pointer, we saw a significant increase in repeated phrases from 1.2% to 2.2%. Coverage was able to bring this down to 1.6%. Coverage tracks which words have already been attended to and copied and prevents copying twice. However, it is still possible to have repetition as our implementation doesn't prevent the character decoder from generating the a word/phrase that has been copied.
- **Delexicalization:** Given low-resource challenge of this project, a strategy drawn from Juraska et al., [4] to improve the generalizability of the model replacing all extracted phrases with the phrase_tag and the POS_tag. Although it shows the best quantitative metrics, the output resembles a template. An example of a delexicalized sequence pair is shown below:

> **Input:** NAME__PROPN, CUISINE__ADJ, DISH_NOUN__NOUN, DISH_NOUN__NOUN
>
> **Target:** NAME__PROPN offers a CUISINE__ADJ menu with a selection of DISH_NOUN__NOUN and DISH_NOUN__NOUN .

- **Beam Search:** In the course of our testing, we also tested beam sizes of 1, 5 and 10.

**Experiment Results:** Below are the results for the test dataset comparing 3323 sequence pairs.

Table 3: Evaluation Metrics: comparing input phrases with output sentences

| Model | Gen Words | Recall (phrase) | Precision (word) | Repeat (word) | Human Good | Human Bad |
|---|---|---|---|---|---|---|
| Vanilla BZ5 | 42967 | 69% | 28% | 1.2% | 27% | 30% |
| Pointer no coverage BZ5 | 41753 | 74% | 32% | 2.2% | 20% | 53% |
| Pointer w coverage BZ5 | 41278 | 74% | 32% | 1.6% | 17% | 40% |
| Pointer w coverage BZ1 | 57584 | 90% | 30% | 2.0% | 13% | 47% |
| Pointer w coverage BZ10 | 38357 | 73% | 34% | 1.5% | 33% | 40% |
| Vanilla BZ5 delexicalized | 50869 | 99% | 47% | N/A | 53% | 10% |

## 4.4 Step 4: Final assessment of text summaries

The final step of passing multiple articles through BERT, generating phrases, clustering and then generating a multi-sentence summary was attempted. However, results were too unsatisfactory to warrant further evaluation and perhaps a bit too ambitious for the time available. Some examples are shown in Appendix C.

# 5 Qualitative Analysis

## 5.1 Extractive Phase - Results Analysis

In the quantitive evaluation of the extractive model, the results were found to exhibit high recall but low precision. When the errors were investigated further, it was found that in many cases of low-precision, the error was often caused by ambiguity in the what should/shouldn't be tagged rather than necessarily the model being wrong. As this is the first time we are tagging this data in this way, we are also learning that the human annotation needs to be better controlled.



In the example 1, the predicted tags included the additional words *"juicy and tender"* as part of the DISH_NOUN. It could work either way and our humans probably tagged it one way in some places and another way in others and created inconsistencies in the labelling.

In the example 2, the predicted tags included the *"live fish"* as a DISH_NOUN whereas the human annotation did not tag those words. This is not a dish that can be ordered in the restaurant so it shouldn't be tagged. The model might be making tagging decisions based too much on the actual words and not enough on the surrounding context.

## 5.2 Generative Phase - Results Analysis

We begin by examining a specific example. The phrases highlighted in yellow from the original text was provided to the encoder-decoder model to generate a sentence.

- **Vanilla model:** The vanilla model had trouble generating the Chef's name given the unusual spellings. Also, it got lost whilst attending to the *"unconventional flavors"* phrase.
- **Pointer with beam size 1:** The pointer with coverage model was able to get the chef's name completely correct as well as replicate the other key phrases but then started to make up additional information. The generated sentence being not much shorter than the original.

- **Pointer with beam size 5:** With the increase in beam size, copying started to fail again in the replication of the Chef's name. The blue phrase was also redundant and repetitive.

- **Pointer with beam size 10:** The Chef's name is still a problem but was, overall, a relatively accurate and well summarized sentence. It is in not purely extractive as we can see some creativity by the use of the word *"offers"* when the original text used the word *"interprets"*.

| | |
|---|---|
| **ORIGINAL** | Chef Kwon Woo Joong interprets traditional Korean cuisine with a decidedly modern flair, using both rare and readily-available seasonal ingredients to create unconventional flavors. |
| **VANILLA** | Chef Kono Woo Joon serves traditional Korean cuisine with modern flair and unop unconners with an exclusive menu. |
| **POINTER BZ1** | Chef Kwon Woo Joong serves traditional Korean cuisine with a modern flair and unconventional flavors that are also a selection of more unfussy flavours. |
| **POINTER BZ5** | Chef Kono Woo Joon serves traditional Korean cuisine with modern flair and unop unconventional flavors to create a selection of more unforgettable flavors. |
| **POINTER BZ10** | Chef Kono Woo Joon offers traditional Korean cuisine with modern flair and unop unconventional flavours. |

**Effect of beam-size:** It's quite interesting to observe how a smaller beam size corresponds to longer sentences and lower precision but also results in better ability to copy (resulting in a better recall rate as shown in the experimental results). The beam size can be used as a dial to exert some control over the abstractive vs. extractive behavior of the model. At lower beam-sizes, the model starts to be too creative and starts to introduce grammatical errors and conjuring erroneous comments.

**Visualization of Attention and Pointer Mechanism:** In the diagram below, the we trace the source of each word generated by our Pointer-Coverage model with beam size of 1. A few notable points:

- The Chef's name had a $P^{att}$ close to 100% and hence the whole name was copied. Without the pointer, we would have *"Kono Woo Joon"* instead of the correct *"Kwon Woo Joong"*.

- The pointer mechanism was also able to save the day by copy *"unconventional"* rather than use *"unop"* generated by the character decoder.

- It is interesting to see that when the model copied over the word *"flavors"*, it discarded *"flavors."* generated by the character decoder, which notably, had a full-stop at the end. This would've been the right place to stop the sentence but because we overwrote the full-stop, the sentence continued with a series of low probability words. Interesting to note that when the model is being "creative" it tends to attend to the commas in between phrases.

- Towards the end of the sentence, it attended strongly on *"unconventional"* again but at this point, the coverage feature kicked in and blocked the copying of the source word, even though the $P^{char_gen}$ of *"unfussy"* was relatively low at 68%.
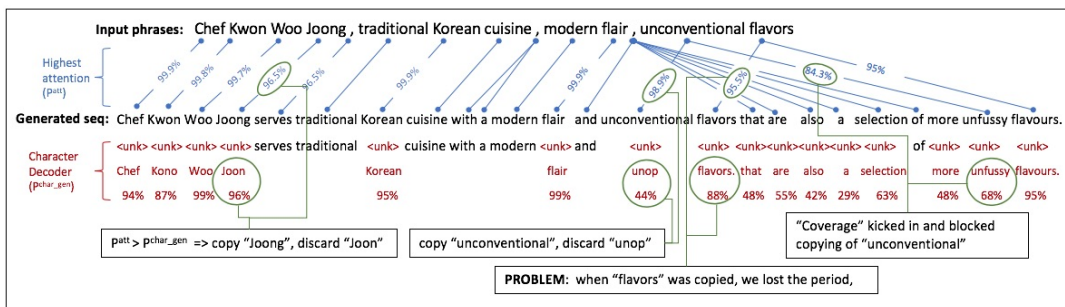


# 6 Conclusion and Future Work

Abstractive text summarization is a challenge and even more so when applied to a new task where no large corpus of document-summary pairs exist for supervised learning.

In the extractive phase of our project, by leveraging on the power of BERT which has been pre-trained on billions of words, we have been able to fine-tune with a very small custom annotated dataset of 1500 sentence with reasonable success. A key learning in the annotation process is that more prescriptive instructions need to be provided to humans to make the annotation more deterministic. Many of the mistakes that the model is making currently is due to ambiguity in the training dataset.

In the generative phase, we were able to obtain 30-50% "good" sequences (as determined by humans) in the sentence-level generation but when we attempted to provide clustered phrases by type extracted from multiple articles to generate a multi-sentence summary, the results were poor.

Firstly, more synthetic data can be generated to train the encoder-decoder without too much investment. This should have a significant improvement on the performance as a conditional language model.

Secondly, in retrospect, perhaps too much time investment was put on the delexicalization experiments. Although it generated the highest quantitative scores, this approach is highly inflexible and takes the power away from what a deep learning model can do with enough data. In essence it ended up learning a lot of templates instead of true abstractive text generation.

Finally, from an architectural standpoint, in only providing the extracted phrases without providing the full text to the encoder-decoder in inference mode, we seem to be throwing a lot of key information that the model needs to fully understand the original text. As a human, it is quite hard to write summaries with only the extracted phrases without the context. For future work, there could be mileage in combining the concepts in this project with the "Key Information Guide Network" (Li et al., 2018 [5]) where both the original text as well as the extracted phrases are provided to the model.

# References

[1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate". In: *arXiv preprint arXiv:1409.0473* (2014).

[2] Sumit Chopra, Michael Auli, and Alexander M Rush. "Abstractive sentence summarization with attentive recurrent neural networks". In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2016, pp. 93–98.

[3] Ondřej Dušek, Jekaterina Novikova, and Verena Rieser. "Findings of the E2E NLG challenge". In: *arXiv preprint arXiv:1810.01170* (2018).

[4] Juraj Juraska et al. "A deep ensemble model with slot alignment for sequence-to-sequence natural language generation". In: *arXiv preprint arXiv:1805.06553* (2018).

[5] Chenliang Li et al. "Guiding generation for abstractive text summarization based on key information guide network". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*. Vol. 2. 2018, pp. 55–60.

[6] Minh-Thang Luong and Christopher D Manning. "Achieving open vocabulary neural machine translation with hybrid word-character models". In: *arXiv preprint arXiv:1604.00788* (2016).

[7] Minh-Thang Luong and Christopher D Manning. "Stanford neural machine translation systems for spoken language domains". In: *Proceedings of the International Workshop on Spoken Language Translation*. 2015, pp. 76–79.

[8] Ramesh Nallapati et al. "Abstractive text summarization using sequence-to-sequence rnns and beyond". In: *arXiv preprint arXiv:1602.06023* (2016).

[9] Alexander M Rush, Sumit Chopra, and Jason Weston. "A neural attention model for abstractive sentence summarization". In: *arXiv preprint arXiv:1509.00685* (2015).

[10] Abigail See, Peter J Liu, and Christopher D Manning. "Get to the point: Summarization with pointer-generator networks". In: *arXiv preprint arXiv:1704.04368* (2017).

[11] Rico Sennrich, Barry Haddow, and Alexandra Birch. "Improving neural machine translation models with monolingual data". In: *arXiv preprint arXiv:1511.06709* (2015).

[12] Ashish Vaswani et al. "Attention is all you need". In: *Advances in Neural Information Processing Systems*. 2017, pp. 5998–6008.

## APPENDIX A

**BERT Architecture - Adapted for Phrase Extraction**

Based on repository: https://github.com/huggingface/pytorch-pretrained-BERT

### BERT for Phrase Extraction
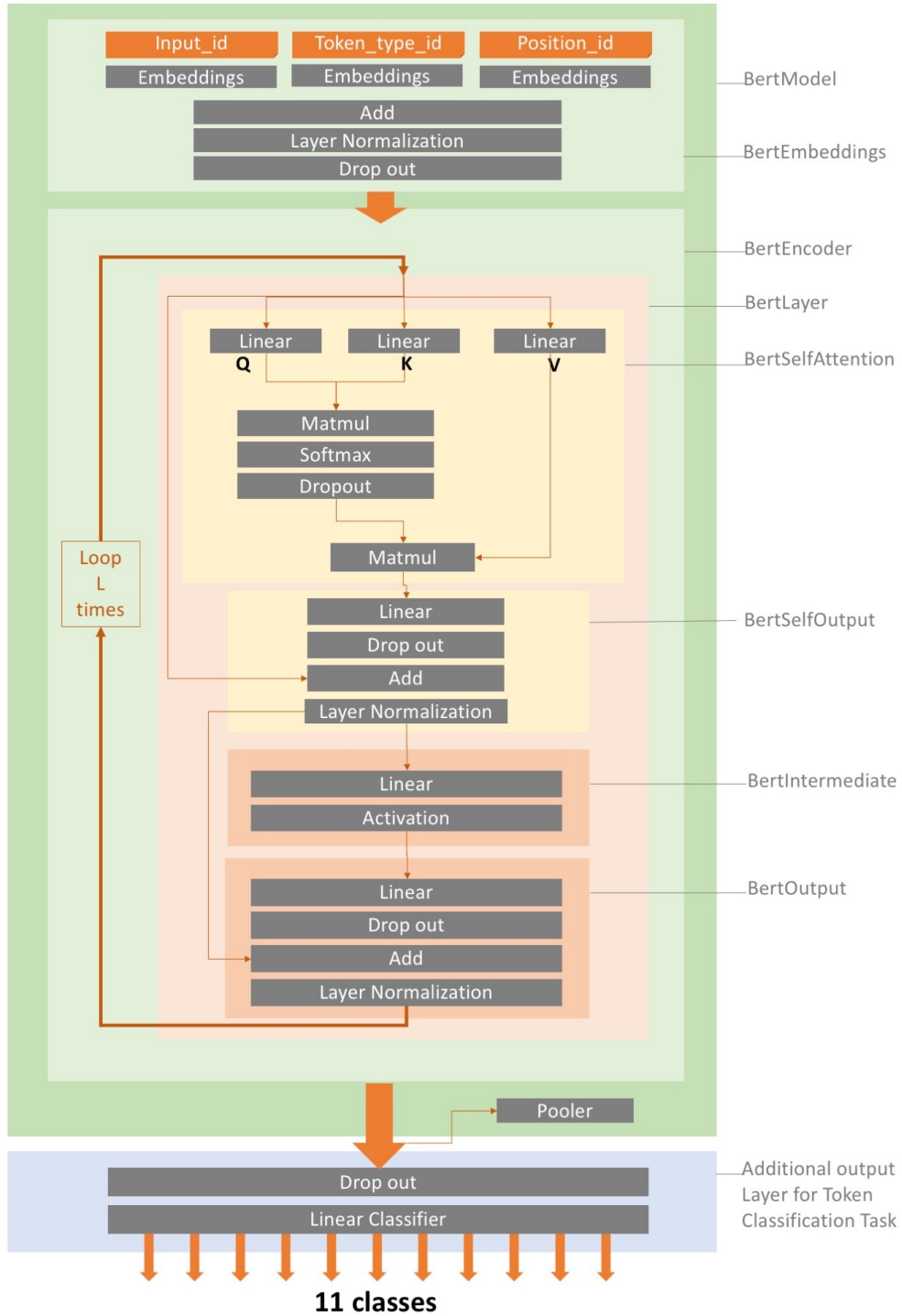
## APPENDIX B

Using the Prodigy tagging tool, $\tilde{2}000$ sentences from the Michelin Guide were tagged by humans with the 11 custom tags.



NAME 1  TYPE 2  CUISINE 3  DECOR_DESC 4  FOOD_DESC 5
DISH_NOUN 6  DRINKS_NOUN 7  CHEF 8  LOCATION 9
HISTORY 10  SERVICE 11

The menu boasts a number of **classic regional dishes** CUISINE including **pike balls with Riesling sauce , linguini and spinach** DISH_NOUN .



NAME 1  TYPE 2  CUISINE 3  DECOR_DESC 4  FOOD_DESC 5
DISH_NOUN 6  DRINKS_NOUN 7  CHEF 8  LOCATION 9
HISTORY 10  SERVICE 11

**Amtskeller** NAME 's **impressive** DECOR_DESC setting - a **wonderful natural stone , barrel - vaulted cellar** DECOR_DESC fitted out in **modern style** DECOR_DESC - is more than matched by the **contemporary** FOOD_DESC food .



NAME 1  TYPE 2  CUISINE 3  DECOR_DESC 4  FOOD_DESC 5
DISH_NOUN 6  DRINKS_NOUN 7  CHEF 8  LOCATION 9
HISTORY 10  SERVICE 11

Enjoy **seasonal classic cuisine** CUISINE including dishes such as **Berlin - style calves ' liver** DISH_NOUN and **salmon with curly kale and chestnut foam** DISH_NOUN .

**APPENDIX C**

**Restaurant Name:** Jade Dragon Restaurant

**Source Text - Michelin Guide Description**

Traditional Chinese art, ebony, crystal, gold and silver converge with modern design to form this stunning and eminently comfortable Cantonese restaurant. Equal thought has gone into the details, such as the striking carved jade chopstick holder. The specialities to look out for are goose grilled over lychee wood, and barbecued Ibérico pork. Seafood is also a highlight, along with herbal soups and recipes based on traditional medicine.

**Source Text - World's 50 Best Restaurants Description:**

In a nutshell: A playground of boutique shops, nightclubs, theatres and casinos, the vast City of Dreams resort in Macao is also home to more than 20 restaurants. The pick of them is this grand Cantonese fine-dining destination that has long wowed critics with its luxurious take on Cantonese cuisine. What to order: Don't leave without trying the Iberico pork barbecued over lychee wood – pluma and collar cuts are both available. Other highlights include hot and sour Kagani crabmeat soup and deep-fried Fine de Claire oysters, while the dim sum menu has proved wildly popular. What's the space like? As lavish as you would expect in the City of Dreams. The palatial restaurant is split into five different areas, each with its own opulent design and colour scheme. Expect plenty of gold, jade, embroidery, thick carpets and beautiful Chinese drawings and ceramics. Other ventures: Kwok Fung has worked across Asia, cooking for the royal families of Thailand, Japan and Nepal, plus politicians and presidents. Did you know? The restaurant has 11 private dining rooms and some are connected directly to the car park by private lifts so famous guests can enter in complete secrecy. Bonus point: Thanks to an exclusive collaboration with the Macau University of Science and Technology, there is a special range of seasonal Chinese herbal soups designed to keep you as healthy as you are satisfied.

**Input Sequences:**

Traditional Chinese art , ebony , crystal , gold and silver , modern design , palatial , opulent design , colour scheme, gold , jade , embroidery , thick carpets , beautiful Chinese drawings , ceramics, private dining rooms goose grilled over lychee wood , barbecued Ibxrico pork, Seafood , herbal soups , traditional, Iberico pork barbecued over lychee wood x pluma , collar cuts, hot and sourxKagani crabmeat soup , deep - fried Fine de Claire oysters , dim sum menu, seasonal Chinese herbal soups comfortable Cantonese restaurant, grand Cantonese fine - dining destination , luxurious take on Cantonese cuisine

**Automated summarization - baseline model:**

Traditional Chinese art with an exposed exposed exposed and exquisite combination of crystal, gold and silver and modern design and a place with a plate of modern design and a colourful scheme gold and ceramical Chinese draws on the ceramical private dining rooms. The restaurant is named after its goose grilled over lychen wood, barbecued Ibxrico pork Seafood and herbal soups with a traditional Iberico pork barbecued over lychee wood x plump, and dim sum menu and dim sum menu. A comfortable Cantonese restaurant grand Cantonese restaurant grand Cantonese dining rooms and a luxurious take on Cantonese cuisine.

**Restaurant Name:** Bo.lan

**Source Text - Michelin Guide Description**

Chef couple Duangporn Songvisava (Bo) and Dylan Jones (lan) have been serving royal Thai dishes to much fanfare for nearly a decade. The heritage of the cuisine is respected, but there is a touch of modernity in the presentation. Real care and effort is put into the ingredients and flavour, along with a zero-carbon goal - choose from three degustation menus that demonstrate the chefs' passion. The attractive and intimate villa is charmingly run.

**Source Text - World's 50 Best Restaurants:**

Then and now: Thai-born chef Duangporn 'Bo' Songvisava and Australian-born chef Dylan 'Lan' Jones met at David Thompson's erstwhile London outpost Nahm. Now, the husband-and-wife team oversees one of the most respected restaurants in Bangkok, and Songvisava was voted Asia's Best Female Chef in 2013. What to expect: Bo.Lan's authentic Thai cooking draws influence from the country's fiery street food and home-cooking traditions as well as its more refined palace-style cuisine. Only the tastiest organic, locally sourced products are used. The menu: In its latest incarnation, the chefs' signature menu 'Bo.Lan balance' is composed of five essential dishes: a Thai salad, a chilli relish, a stir-fry, a curry and a soup – although more dishes may be added to enhance the experience at the chefs' discretion. What's it like inside? Having moved location a few years ago, it is now one of the most glamorous restaurants in the city. The space has a contemporary feel with plenty of beautiful untreated wood, while maintaining a rich, traditional Thai ambience. Who to dine with: Take someone who shares an adventurous and discerning appetite as dinner is only served in set menus in a sharing style. At lunch, there is also an à la carte menu. What's cool: Chefs Songvisava and Jones are committed to keeping a zero carbon footprint. To achieve this, they have built a vegetable garden, a water filtering system and waste recycling scheme.
Chef coule Duongan Sonannese and royal Thai dishes are served in this former neatly a decadent. In this modern restaurant in an intimate villa with a charming run contemporary feel and a beautiful enchanting contemporary feel to the rich and traditional Thai ambience. It serves authentic Thai cooking with fiery street food, home-cooked traditional and home-cooked traditional dishes from the seasons of textures and flavours to delicate spirits. Lande balance arrives with Thai salad and chilli relish, stir-fried and sour.

**Input Sequences:**

Chef couple Duangporn Songvisava , royal Thai dishes , for nearly a decade
modernity attractive , intimate villa , charmingly run, contemporary feel , beautiful untreated wood , rich , traditional Thai ambience
Bo . Lanxs authentic Thai cooking , fiery street food , home - cooking traditions , palace - style cuisine, artful , original , creative , of textures , flavours , delicate spicing
Lan balancex , Thai salad , chilli relish , stir - fry , curry , soup

**Automated summarization - baseline model:**

Chef coule Duongan Sonannese and royal Thai dishes are served in this former neatly a decadent. In this modern restaurant in an intimate villa with a charming run contemporary feel and a beautiful enchanting contemporary feel to the rich and traditional Thai ambience. It serves authentic Thai cooking with fiery street food, home-cooked traditional and home-cooked traditional dishes from the seasons of textures and flavours to delicate spirits. Lande balance arrives with Thai salad and chilli relish, stir-fried and sour.

**Restaurant Name:** Gaggan

**Source Text - Michelin Guide Description**

Chef-owner Gaggan Anand takes Indian cuisine to a rarely seen level and one that is pure alchemy. His artful dishes are original and creative, with a wonderful blend of textures, flavours, and delicate spicing. To see the dishes unfold before you, ask for the counter. It's a feast for the senses as dishes are delivered at a rapid pace by the passionate staff. With 30 chefs in the kitchen there's plenty of manpower on hand to deliver a memorable experience.

**Source Text - World's 50 Best Restaurants:**

Chief reason to visit: Gaggan Anand's fertile imagination and culinary wit are played out in a lengthy but always fun tasting menu which balances the soulfulness of Indian street food with hyper-modern cooking techniques to unrivalled effect. A word on the chef: Kolkata-born Anand is something of a superstar in the food world, as much for his effervescent personality as his boundary-pushing dishes. After a life-changing stint in the kitchen at El Bulli, the chef opened his eponymous Bangkok restaurant in 2010, where he has reimagined the notion of 'progressive Indian cuisine'. He has since opened burger and beer restaurant Meatlicious and is launching natural wine bar Wet and tofu restaurant Mihara Tofuten in 2018. On the tasting: The 25-course menu, written entirely in emojis, starts off with a series of rapid-fire snacks including the spherified Yoghurt Explosion, and proceeds through a long sequence of courses including Indian sushi, charcoal prawn Amritsari and uni ice cream in a miniature cone. At the end of the meal, guests are presented with a version of a menu that contains the names of the dishes. Pick a highlight: Anand's latest signature 'Lick it up' is served in time with the eponymous song by legendary rock band Kiss. Made with green peas, fenugreek, mushroom and tomato, the dish is designed to be 'licked up' directly from the plate – the flavours are layered to touch the tongue exactly where the different taste receptors are placed. What's the vibe? The restaurant is set in an elegant whitewashed and recently renovated colonial house, tucked away from the mayhem of the city's busy streets, with a number of different dining spaces. Service is relaxed – despite the occasional theatrics created by the use of liquid nitrogen – and the cocktail list is particularly strong. Try to bag the table adjacent to a giant window into the kitchen to watch the multi-national brigade show off their alchemy, or sit upstairs at the chef's table in the brand new R&D lab. Is it really the Best Restaurant in Asia? According to the 300-plus voters, yes! Gaggan is boundary-pushing but never too serious: a restaurant where fun dining is taken to a new level, but not at the expense of taste or technique. Bonus point: Gaggan starred in his own episode of the hit Netflix series Chef's Table in 2016.

**Input Sequences:**

Gaggan Anandxs , soulfulness , Indian street food , hyper - modern , unrivalled, xprogressive Indian cuisinex effervescent personality
El Bulli , Bangkok restaurant in 2010 , burger and beer restaurant Meatlicious , wine bar Wet and tofu restaurant Mihara Tofuten in
emojis , spherified Yoghurt Explosion , Indian sushi , charcoal prawn Amritsari , uni ice cream in a miniature cone Anandxs
green peas , fenugreek , mushroom and tomato
elegant whitewashed and recently renovated colonial house , busy, giant window , chef 's table , R&D lab cocktail list

**Automated summarization - baseline model:**

Gango Anna serves soulfish and Indian street food in a home-grown and uncomplicated awards Indian cuisine. The extensive menu features excellent value for money. El Bulli has been converted into a fine Bangkok restaurant in 2015. A burger and beer restaurant Mediterranean is a fine wine bar Wet and tofu restaurant. For dessert, studdends and specialises in stuffed Yoguet Exploring Indian sushi and chili ice cream in a minimalist conservatory. Andine and the chef consists of the chef's own contemporary country starters. Try green peas, fennel, mushroom and tomato. An elegant whitewashed and recently renovated colonial house with a busy giant window and a busy giant window and a classic Ramage lake. The cocktail list is also worth a try.

**Restaurant Name:** The Chairman

**Source Text - Michelin Guide Description**

The Chairman looks to small suppliers and local fishermen for its ingredients and much of the produce used is also organic. Showing respect for the provenance of ingredients, and using them in homemade sauces and flavoursome dishes such as steamed crab with aged Shaoxing, crispy chicken stuffed with shrimp paste and almond sweet soup, has attracted a loyal following. The restaurant is divided into four different sections and service is pleasant and reassuringly experienced.

**Source Text - World's 50 Best Restaurants:**

The location: Hunting down this two-storey restaurant tucked in the corner of Kau U Fong on the undulating laneway of Central is a feat but let's just say diners will be richly rewarded. What to order: The perennial favourite among foodies is chef Kwok Keung Tung's steamed flowery crab in aged Shaoxing wine and chicken oil with flat rice noodles. Their ethos: The Chairman does more than pay lip service to its locavore intention. Its chickens and pigs are sourced locally from New Territories, the seafood is purchased from local fishermen who brave the South China Sea for their daily catch of live shrimps and fish; the restaurant even owns a small organic farm in Sheung Shui where they plant vegetables that, in future, will cater to a portion of the restaurant's needs. The owner: Danny Yip is a low-profile food cognoscenti who much prefers to let the restaurant and its cuisine do the talking. He used to own restaurants in Canberra, Australia, before he returned to Hong Kong to open The Chairman in 2009. Bonus point: This year, The Chairman wins the Highest Climber Award, rising 25 places in the list – tied only with Mume in Taipei.

**Input Sequences:**

homemade sauces , steamed crab with aged Shaoxing , crispy chicken stuffed with shrimp paste , almond sweet soup, chef Kwok Keung Tungxs steamed flowery crab in aged Shaoxing wine and chicken oil with flat rice noodles, chickens , shrimps , Sheung Shui pleasant , reassuringly experienced
two - storey , corner of Kau U Fong , undulating laneway of Central

**Automated summarization - baseline model:**

The menu offers a range of homemade sauces, steamed crab with aged Shanghai, crispy chicken stuffed with aged Shanghai, crispy chicken stuffed with shrimp paste, almond sweet soup chef Kwon Keung Tunals steamed flower crab in aged Shanghai wine and chicken oil with flat rice noodles chicken and shrimps and Sheffield Shui and the selection of specials on offer in the restaurant. The atmosphere is pleasant and reasonably excellent. A two-storey restaurant in the corner of Kau U Fog, this restaurant is housed in a unusual landmark of Central in the evening.

**Restaurant Name:** Mingles

**Source Text - Michelin Guide Description**

> Since 2014, this contemporary fine dining establishment has captured the imagination of even the most discerning palates with Chef Kang Mingoo's new and bold creations that pay respect to Korea's culinary heritage. Traditional fermented condiments and vinegar-'jang' and 'cho'-play integral roles in the dishes, even dessert. Traditional liquor pairing is offered in addition to wine pairing.

**Source Text - World's 50 Best Restaurants:**

> Why it's worth a visit: Since its opening in April 2014 in Seoul's buzzy Cheongdam-dong district, Mingles has been wowing local and international diners with its fusion of eastern and western cuisine. So much so, that the restaurant was the Highest New Entry to Asia's 50 Best Restaurants 2016 and has now been voted Best Restaurant in Korea for the third consecutive year. Who's behind the pass? Young Korean chef Mingoo Kang trained under Martin Berasategui in San Sebastian, Spain, and later enjoyed stints at Nobu in Miami and the Bahamas, where he took learnings that he would later apply to Korean food once back on home soil. Key ingredients: The seasonally changing menu is divided into sections like 'grains,' 'vegetable' and 'fish,' with jang and cho – Korean traditional fermented sauce and vinegar – playing an integral part in the food. Main courses include bansang, a Korean meal with rice, soup, kimchi and 'Jang' sauce, and guests can enjoy the zucchini seon – a special dish first conceived as a vegan one-off when a famous monk visited the restaurant – with green pumpkin and zucchini cooked with anchovy broth and salted shrimp. The dish takes over two days of preparation. The vibe: Mingles' stripped-back basement dining room wouldn't look out of place in London or New York with its blonde-wood furniture and smooth stone walls, and the casual atmosphere makes it the go-to restaurant for hip Seoulsters. Bonus point: Kang was among the 1,500 runners who carried the Olympic torch 2,018km through Korea's main cities and to the PyeongChang Olympic Stadium for the opening ceremony of the 2018 Winter Olympics.

**Input Sequences:**

> Since 2014 , contemporary fine dining , Chef Kang Mingoo , new , bold creations, Since its opening in April 2014 , Seoulxs buzzy Cheongdam - dong district , Mingles , eastern and western cuisine Traditional fermented condiments , vinegar-'jang Traditional liquor pairing , wine pairing

**Automated summarization - baseline model:**

> Since 2014, the contemporary fine dining experience at Chef Kang Mingli is a simple new and bold creative Since its open in Arizzi 2014 in Mingles, eastern and west cuisine. Traditional fermented contemporary features and the menu is a starter of the chef's signature dishes. Traditional liquor pairing is accompanied by an extensive wine pairing and an extensive wine list.