1 Dataset Details

CNN/DailyMail Dataset CNN/DailyMail dataset (Hermann et al., 2015; Nallapati et al., 2016) is a large collection of online news articles and their multi-sentence summaries. We use the original, non-anonymized version of the dataset provided by See et al. (2017). Overall, the dataset has 287,226 training pairs, 13,368 validation pairs and, 11,490 test pairs. On an average, a source document has 781 tokens and a target summary has 56 tokens.

Gigaword Corpus Gigaword is based on a large collection of news articles, where the article’s first sentence is considered as the input document and the headline of the article as output summary. We use the annotated corpus provided by Rush et al. (2015). It has around 3.8 million training samples. For validation, we use 2,000 samples and for test evaluation we use the standard test set provided by Rush et al. (2015). Following previous work, we keep our vocabulary size to 50,000 frequent words.

DUC Corpus We use the DUC-20021 document summarization dataset for checking our model’s generalizability capabilities. DUC-2002 corpus consists of 567 documents with one or two human annotated reference summaries. We also tried beam retuning using DUC-20032 as a validation set, which consists of 624 documents with single human annotated reference summaries.

SNLI corpus We use the Stanford Natural Language Inference (SNLI) corpus (Bowman et al., 2015) for our entailment generation task. Following Pasunuru and Bansal (2017), we use the same re-splits provided by them to ensure a zero train-test overlap and multi-reference setup. This dataset has a total of 145,822 unique premise pairs out of 190,113 pairs, which are used for training, and the rest of them are divided equally into validation and test sets.

SQuAD Dataset We use Stanford Question Answering Dataset (SQuAD) for our question generation task (Rajpurkar et al., 2016). In SQuAD dataset, given the comprehension and question, the task is to predict the answer span in the comprehension. However, in our question generation task, we extract the sentence from the comprehension containing the answer span and create a sentence-question pair similar to Du et al. (2017). The dataset has around 100K sentence-question pairs from 536 articles.

2 Training Details

The following training details are common across all models and datasets. We use LSTM-RNN in our sequence models with hidden state size of 256 dimension. We use 128 dimension word embedding representations. We do not use dropout or any other regularization techniques, but we clip the gradient to allow a maximum gradient norm value of 2.0. We use Adam optimizer (Kingma and Ba, 2015) with a learning rate of 0.001. Also, we share the word embeddings representation of both encoder and decoder in our models. All our tuning decisions (including soft/hard and layer-specific sharing decisions) were made on the appropriate validation/development set.

CNN/DailyMail: For all the models involving CNN/DailyMail dataset, we use a maximum encoder RNN step size of 400 and a maximum decoder RNN step size of 100. We use a mini-
John Hughes has revealed how he came within a heartbeat of stepping down from his job at Inverness as the Josh Meekings controversy went into overdrive this week. The Caley Thistle boss says he felt so repulsed by the gut-wrenching predicament being endured by his young defender - before he was dramatically cleared - that he was ready to walk away from his post and the games he loves, just weeks before an historic Scottish Cup final date. Keen cyclist Hughes set off on a lonely bike ride after hearing Meekings had been cited for the handball missed by officials in the semi-final against Celtic, and admits his head was in a spin over an affair that has dominated the back-page headlines since last Sunday. Inverness defender Josh Meekings will be allowed to appear in Scottish Cup final after his ban was dismissed. Only messages of support awaiting him on his return from footballing friends brought him back from the brink of quitting. Hughes, who lives in the Black Isle just north of Inverness, said: ‘I came in here this morning after a day off. I turned my phone off and was away myself, away out on the bike with plenty of thinking time: a great freedom of mind. ‘I was that sick of what has been going on in Scottish football I was seriously contemplating my own future. I’m serious when I say that. ‘I had just had it up to here and was ready to just give it up. If it was n’t for what happened when i turned my phone back on, with the phone calls and texts I received from people I really value in football, that my spirits picked up again. ‘The calls and texts came in from all over the place, from some of the highest levels across the game. ‘I’ve had phone calls that have really got me back on my feet. ‘I would n’t like to name them all, but there were a lot of good people and a good few close friends in the football fraternity. Meekings was not sent off and no penalty was given as Inverness went on to beat Celtic 3-2 after extra-time. ‘They were saying: “You need to lead from the front, you need to fight it.” That restored and galvanised that focus and drive in me. And, if that was how I was feeling, how was the boy Josh Meekings feeling? It should never have come to this. ‘Meekings was cleared to play in the final by the judicial panel yesterday, but Hughes insists this ‘unprecedented’ SFA wrangle must be the catalyst for change in Scottish football’s governance. Although those who sit on the panel are drawn from many walks of life, ranging from former players and coaches to ex-refs and members of the legal profession, Hughes said he wants ‘real football people’ ‘drafted in instead of the ‘suits’ he claims lack understanding of the nuances and spirit of the professional game. And he seemed to point a thinly-veiled finger of accusation at SFA chief executive Stewart Regan by alleging that compliance officer Tony McGlenan was a mere ‘patsy’ in the process. (...)

Inverness defender Josh Meekings has won appeal against one-match ban. The 22-year-old was offered one-game suspension following incident. However, an independent judicial panel tribunal overturned decision. Inverness reached the Scottish Cup final with 3-2 win over Celtic. See et al. (2017): Josh Meekings has been cleared to play in the Scottish Cup final. The Englishman admitted he was fortunate not to have conceded a penalty and been sent off by [referee Steven McLean] for stopping Leigh Griffiths net-bound effort on his goal-line. Meekings was not sent off and no penalty was given as Inverness went on to beat Celtic 3-2.

Our Baseline: Josh Meekings cleared to play in the Scottish cup final on [May 30]. Inverness defender Josh Meekings will be allowed to appear in Scottish cup final. Meekings was not sent off and no penalty was given as Inverness went on to beat Celtic 3-2.

Multi-task: Josh Meekings has been cleared to play in the Scottish cup final. Inverness defender Josh Meekings will be allowed to appear in Scottish cup final after his ban was dismissed. Inverness went on to beat Celtic 3-2 after extra-time.

Figure 1: Example showing summaries generated by See et al. (2017), our baseline, and 2-way multi-task model with summarization and entailment generation. Boxed-red highlighted words/phrases are not present in the input source document in any paraphrasing form. As shown, both See et al. (2017) and the baseline generate extraneous information that is not entailed by the source documents ("referee Steven McLean" and "May 30"), but our multi-task model avoids such unrelated information to generate summaries that logically follow from the source document.
bending and rising in spectacular fashion, these stunning pictures capture the paddy fields of south east asia and the arduous life of the farmers who cultivate them. in a photo album that spans over china, thailand, vietnam, laos and cambodia, extraordinary images portray the crop’s full cycle from the primitive sowing of seeds to the distribution of millions of tonnes for consumption. the pictures were taken by professional photographer scott gable, 39, who spent four months travelling across the region documenting the labour and threadbare equipment used to harvest the carbohydrate-rich food. scroll down for video. majestic: a farmer wades through the mud with a stick as late morning rain falls on top of dragonsbone terraces in longsheng county, china. rice is a staple food for more than one-half the world’s population, but for many consumers, its origin remains somewhat of a mystery. the crop accounts for one fifth of all calories consumed by humans and 87 per cent of it is produced in asia. it is also the thirstiest crop there is - according to the un, farmers need at least 2,000 litres of water to make one kilogram of rice. mr gable said he was determined to capture every stage of production with his rice project - from the planting to the harvesting all the way down to the shipping of the food. after acquiring some contacts from experts at cornell university in new york and conducting his own research, he left for china last may and spent the next four months traveling. he said: ‘the images were taken over a four month period from april to july last year across asia. i visited china, thailand, vietnam, laos and cambodia as part of my rice project. video courtesy of www.scottgable.com. breathtaking: a paddy field worker toils on the beautiful landscape of dragonsbone terraces in longsheng county, china. farmers’ procession: a rice planting festival parade takes place near the village of pingan in guangxi province, china. ‘the project is one part of a larger three part project on global food staples - rice, corn and wheat. i am currently in the process of shooting the corn segment. ‘the industrialisation of our food and mono-culture food staples have interested me for some time so that’s probably what inspired me to do this project. ‘i shot the whole project using a canon slr and gopros. the actual shooting took four months and then post production took another four more months. ‘the reaction to my work has been incredibly positive - i was able to secure a solo gallery show and create quite a bit of interest online which has been great. ‘family crop: a hani woman in traditional clothing sits on top of her family’s rice store in yunnan province, china. arduous labour: employees of taiwan’s state-run rice experimental station are pictured beating rice husks by hand as the sun shines on them. mr gable spent months learning mandarin chinese in preparation for his trip, but the language barrier was still his greatest challenge. (...)
Celtic have written to the Scottish Football Association in order to gain an 'understanding' of the refereeing decisions during their Scottish Cup semi-final defeat by Inverness on Sunday. The hoops were left outraged by referee Steven McLean's failure to award a penalty or red card for a clear handball in the box by Josh Meekings to deny Leigh Griffiths' goal-bound shot during the first-half. Caley Thistle went on to win the game 3-2 after extra-time and denied Rory Delia's men the chance to secure a domestic treble this season. Celtic striker Leigh Griffiths has a goal-bound shot blocked by the outstretched arm of Josh Meekings. Celtic's Adam Matthews (right) slides in with a strong challenge on Nick Ross in the Scottish Cup semi-final. 'Given the level of reaction from our supporters and across football, we are duty bound to seek an understanding of what actually happened,' Celtic said in a statement. They added, 'we have not been given any other specific explanation so far and this is simply to understand the circumstances of what went on and why such an obvious error was made.' However, the Parkhead outfit made a point of congratulating their opponents, who have reached the first-ever Scottish Cup final in their history, describing Caley as a 'fantastic club' and saying 'reaching the final is a great achievement.' Celtic had taken the lead in the semi-final through defender Virgil van Dijk's curling free-kick on 18 minutes, but were unable to double that lead thanks to the Meekings controversy. It allowed Inverness a route back into the game and Celtic had goalkeeper Craig Gordon sent off after the restart for scything down Marley Watkins in the area. Greg Tansey duly converted the resulting penalty. Edward Ofere then put Caley Thistle ahead, only for John Guidetti to draw level for the bhoys. With the game seemingly heading for penalties, David Raven scored the winner on 117 minutes, breaking thousands of Celtic hearts. Celtic captain Scott Brown (left) protests to referee Steven McLean but the handball goes unpunished. Griffiths shows off his acrobatic skills during Celtic's eventual surprise defeat by Inverness. Celtic pair Aleksandar Tonev (left) and John Guidetti look dejected as their hopes of a domestic treble end.

Ground-truth: Celtic were defeated 3-2 after extra-time in the Scottish Cup semi-final. Leigh Griffiths had a goal-bound shot blocked by a clear handball, however, no action was taken against offender Josh Meekings. The hoops have written the SFA for an 'understanding' of the decision.

See et al. (2017): John Hartson was once on the end of a major Hampden injustice while playing for Celtic. But he can not see any point in his old club writing to the Scottish Football Association over the latest controversy at the National Stadium. Hartson had a goal wrongly disallowed for offside while Celtic were leading 1-0 at the time but went on to lose 3-2.

Our Baseline: John Hartson scored the late winner in 3-2 win against Celtic. Celtic were leading 1-0 at the time but went on to lose 3-2. Some fans have questioned how referee Steven McLean and additional assistant Alan Muir could have missed the infringement.

Multi-task: Celtic have written to the Scottish Football Association in order to gain an 'understanding' of the refereeing decisions. The hoops were left outraged by referee Steven McLean's failure to award a penalty or red card for a clear handball in the box by Josh Meekings. Celtic striker Leigh Griffiths has a goal-bound shot blocked by the outstretched arm of Josh Meekings.

Figure 3: Example of summaries generated by See et al. (2017), our baseline, and 3-way multi-task model with summarization and both entailment generation and question generation. The boxed-red highlighted words/phrases are not present in the input source document in any paraphrasing form. All the unboxed-green highlighted words/phrases correspond to the salient information. See detailed discussion in Fig. 1 and Fig. 2 above. As shown, the outputs from See et al. (2017) and the baseline both include non-entailed words/phrases (e.g. “John Hartson”), as well as they missed salient information (“hoops”, “Josh Meekings”, “Leigh Griffiths”) in their output summaries. Our multi-task model, however, manages to accomplish both, i.e., cover more salient information and also avoid unrelated information.
batch size of 16. We initialize the LSTM-RNNs with uniform random initialization in the range \([-0.02, 0.02]\). We set \(\lambda\) to 1.0 in the joint cross-entropy and coverage loss. Also, we only add coverage to the converged model with attention and pointer mechanism, and make the learning rate from 0.001 to 0.0001. During multi-task learning, we use coverage mechanism for primary (CNN/DailyMail summarization) task but not for auxiliary tasks (because they do not have traditional redundancy issues). The penalty coefficient \(\gamma\) for soft-sharing is set to \(5 \times 10^{-5}\) and \(1 \times 10^{-5}\) for 2-way and 3-way multi-task models respectively (the range of the penalty value is intuitively chosen such that we balance the cross-entropy and regularization losses). In inference time, we use a beam search size of 4, following previous work (See et al., 2017).

**Gigaword:** For all the models involving Gigaword dataset, we use a maximum encoder RNN step size of 50 and a maximum decoder RNN step size of 50. We use a mini-batch size of 256. We initialize the LSTM-RNNs with uniform random initialization in the range \([-0.01, 0.01]\). We do not use coverage mechanism to our Gigaword models. Also, we set our beam search size to 5, following previous work (Nallapati et al., 2016).

**DUC:** For the CNN/DM to DUC domain-transfer experiments where we allow the beam sizes of all models to be individually re-tuned on DUC-2003, the chosen tuned beam values are 10, 4, 3 for the multi-task model, baseline, and See et al. (2017), respectively.

### 2.1 Multi-Task Learning Details

**Multi-Task Learning with Question Generation** Two important hyperparameters tuned are the mixing ratio between summarization and entailment generation, as well as the soft-sharing coefficient. Here, we choose the mixing ratios 3:2 between CNN/DailyMail and SQuAD, 100:1 between Gigaword and SQuAD. Intuitively, these mixing ratios are close to the ratio of their dataset sizes. We set the soft-sharing coefficient \(\gamma\) to \(5 \times 10^{-5}\) and \(1 \times 10^{-5}\) for CNN/DailyMail and Gigaword, resp.

**Multi-Task Learning with Entailment Generation** Here, we choose the mixing ratios 3:2 between CNN/DailyMail and SNLI, 20:1 between Gigaword and SNLI. We again set the soft-sharing coefficient \(\gamma\) to \(5 \times 10^{-5}\) and \(1 \times 10^{-5}\) for CNN/DailyMail and Gigaword, resp.

---

**Figure 4:** Output examples from our entailment generation model.

**Input:** The college of science was established at the university in 1865 by president father Patrick Dillon.

**Question:** In what year was the college of science established?

**Entailment:** People walk down the street.

**Premise:** People walk down a paved street that has red lanterns hung from the buildings.

**Entailment:** People walk down the street.

**Figure 5:** Output examples from our question generation model.

**Input:** Notable athletes include swimmer sharron Davies, diver Tom Daley, dancer Wayne Sleep, and footballer Trevor Francis.

**Question:** What is the occupation of Trevor Francis?

**Entailment:** A young woman strikes a man with her feet.

**Premise:** A young woman on a boat in a light colored bikini kicks a man wearing a straw cowboy hat.

**Entailment:** A young woman strikes a man with her feet.

---

**References**

Bowman, Samuel R, Angeli, Gabor, Potts, Christopher, and Manning, Christopher D. 2015. A large annotated corpus for learning natural language inference. In EMNLP.

Du, Xinya, Shao, Junru, and Cardie, Claire. 2017. Learning to ask: Neural question generation for reading comprehension. In ACL.


Kingma, Diederik, and Ba, Jimmy. 2015. Adam: A method for stochastic optimization. In ICLR.


Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. 2016. Squad: 100,000+ questions for machine comprehension of text. In EMNLP.

Alexander M Rush, Sumit Chopra, and Jason Weston. 2015. A neural attention model for abstractive sentence summarization. In EMNLP.