

Summary of references

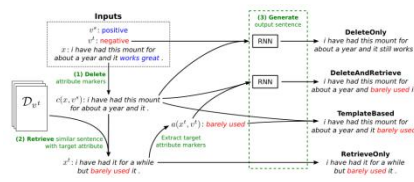
Stanza:

<https://stanfordnlp.github.io/stanza/depparse.html>

Pytextrank:

<https://derwen.ai/docs/ptr/overview/>

Title	Dataset	Method	Evaluation matrix																																																															
《Delete, Retrieve, Generate: A Simple Approach to Sentiment and Style Transfer》	1.YELP	Baseline models:	<table><tr><th></th><th colspan="2">YELP</th><th colspan="2">CAPTIONS</th><th colspan="2">AMAZON</th></tr><tr><th></th><th>Classifier</th><th>BLEU</th><th>Classifier</th><th>BLEU</th><th>Classifier</th><th>BLEU</th></tr><tr><td>CROSSALIGNED</td><td>73.7%</td><td>3.1</td><td>74.3%</td><td>0.1</td><td>74.1%</td><td>0.4</td></tr><tr><td>STYLEEMBEDDING</td><td>8.7%</td><td>11.8</td><td>54.7%</td><td>6.7</td><td>43.3%</td><td>10.0</td></tr><tr><td>MULTIDECODER</td><td>47.0%</td><td>7.1</td><td>68.3%</td><td>4.6</td><td>68.3%</td><td>5.0</td></tr><tr><td></td><td>81.7%</td><td>11.8</td><td>92.5%</td><td>17.1</td><td>68.7%</td><td>27.1</td></tr><tr><td></td><td>95.4%</td><td>0.4</td><td>95.5%</td><td>0.7</td><td>78.3%</td><td>0.9</td></tr><tr><td></td><td>85.7%</td><td>7.5</td><td>83.0%</td><td>9.0</td><td>45.6%</td><td>24.6</td></tr><tr><td></td><td>88.7%</td><td>8.4</td><td>96.8%</td><td>7.3</td><td>48.0%</td><td>22.8</td></tr></table>		YELP		CAPTIONS		AMAZON			Classifier	BLEU	Classifier	BLEU	Classifier	BLEU	CROSSALIGNED	73.7%	3.1	74.3%	0.1	74.1%	0.4	STYLEEMBEDDING	8.7%	11.8	54.7%	6.7	43.3%	10.0	MULTIDECODER	47.0%	7.1	68.3%	4.6	68.3%	5.0		81.7%	11.8	92.5%	17.1	68.7%	27.1		95.4%	0.4	95.5%	0.7	78.3%	0.9		85.7%	7.5	83.0%	9.0	45.6%	24.6		88.7%	8.4	96.8%	7.3	48.0%	22.8
		YELP		CAPTIONS		AMAZON																																																												
		Classifier		BLEU	Classifier	BLEU	Classifier	BLEU																																																										
	CROSSALIGNED	73.7%		3.1	74.3%	0.1	74.1%	0.4																																																										
	STYLEEMBEDDING	8.7%		11.8	54.7%	6.7	43.3%	10.0																																																										
MULTIDECODER	47.0%	7.1	68.3%	4.6	68.3%	5.0																																																												
	81.7%	11.8	92.5%	17.1	68.7%	27.1																																																												
	95.4%	0.4	95.5%	0.7	78.3%	0.9																																																												
	85.7%	7.5	83.0%	9.0	45.6%	24.6																																																												
	88.7%	8.4	96.8%	7.3	48.0%	22.8																																																												
2.Reviews on Amazon		RETRIEVEONLY: Returns, word by word, sentences retrieved that are similar to the original sentence with the deleted attribute. Guaranteed to produce a grammatically correct sentence with the target attributes, but its content may differ from the original sentence.																																																																
3.Image captions		TEMPLATEBASED: Replace the deleted attribute in the original sentence with the target attribute																																																																
		Neural modes:																																																																
	<table><tr><th>Dataset</th><th>Attributes</th><th>Train</th><th>Dev</th><th>Test</th></tr><tr><td rowspan="2">YELP</td><td>Positive</td><td>270K</td><td>2000</td><td>500</td></tr><tr><td>Negative</td><td>180K</td><td>2000</td><td>500</td></tr><tr><td rowspan="3">CAPTIONS</td><td>Romantic</td><td>6000</td><td>300</td><td>0</td></tr><tr><td>Humorous</td><td>6000</td><td>300</td><td>0</td></tr><tr><td>Factual</td><td>0</td><td>0</td><td>300</td></tr><tr><td rowspan="2">AMAZON</td><td>Positive</td><td>277K</td><td>985</td><td>500</td></tr><tr><td>Negative</td><td>278K</td><td>1015</td><td>500</td></tr></table>	Dataset	Attributes	Train	Dev	Test	YELP	Positive	270K	2000	500	Negative	180K	2000	500	CAPTIONS	Romantic	6000	300	0	Humorous	6000	300	0	Factual	0	0	300	AMAZON	Positive	277K	985	500	Negative	278K	1015	500	DELETEONLY: An RNN is used to embed the sentences with the attributes removed into a vector. It then connects the final hidden state to the learned embedding of the target attribute and feeds it to the RNN decoder to generate the final sentence. The decoder attempts to generate words that represent the source content and target attributes while maintaining fluency.																												
Dataset	Attributes	Train	Dev	Test																																																														
YELP	Positive	270K	2000	500																																																														
	Negative	180K	2000	500																																																														
CAPTIONS	Romantic	6000	300	0																																																														
	Humorous	6000	300	0																																																														
	Factual	0	0	300																																																														
AMAZON	Positive	277K	985	500																																																														
	Negative	278K	1015	500																																																														
		DELETEANDRETRIEVE: Similar to DELETEONLY, but uses the attributes of the retrieved sentence rather than the target attributes. As with DELETEONLY, it encodes the sentence with the original deleted attribute using an RNN. The RNN decoder uses the splicing and content embedding of																																																																

		<p>this vector to generate the final sentence.</p> 																																																																																																																																																																																																																																																																					
Text Style Transfer: A Review and Experimental Evaluation	<div>1. Yelp a corpus of restaurant reviews</div> <div>2. Amazon product review datasets</div> <div>3. IMDb movie reviews datasets</div>	<p>Creating a taxonomy to organize the TST models, and provide a comprehensive summary of the state of the art. We review the existing evaluation methodologies for TST tasks, and conduct a large-scale reproducibility study where we experimentally benchmark 19 state-of-the-art TST algorithms on two publicly available datasets. Finally, we expand on current trends and provide new perspectives on the new and exciting developments in the TST field.</p>	<p>Table 10: Example outputs Yelp datasets.</p> <table><tr><th>Model</th><th>From negative to positive (Yelp)</th><th>From positive to negative (Yelp)</th></tr><tr><td>Source</td><td>Food was cold and lacking of flavors.</td><td>Excellent service, and food, and nice atmosphere.</td></tr><tr><td>DeletionOnly</td><td>They were great, with a wide variety of flavors.</td><td>I would give zero stars, when I could.</td></tr><tr><td>Template</td><td>Favorite of flavors and delicious.</td><td>Excellent service, great food, and nice atmosphere.</td></tr><tr><td>DeletionRetrieval</td><td>Of a large list of flavors.</td><td>A waste of time and money.</td></tr><tr><td>B-GST</td><td>Food was bland and of no flavors.</td><td>Great service, great food, and great atmosphere.</td></tr><tr><td>G-GST</td><td>The food was cold and lacking of flavors.</td><td>The service, nice food, and atmosphere.</td></tr><tr><td>PTD</td><td>Food was fresh and delicious of flavors.</td><td>Horrible service, horrible food, and horrible atmosphere.</td></tr><tr><td>SMILE</td><td></td><td>Slow service, good food, and nice atmosphere.</td></tr><tr><td>DRIST</td><td>Food was delicious and fresh.</td><td>Poor service, poor food, and poor atmosphere.</td></tr><tr><td>RET</td><td>Food is delicious, and the staff.</td><td>Service was ok but the food was not hot.</td></tr><tr><td>CAAE</td><td>Food was fresh and good.</td><td>However, food was good, and poor, poor atmosphere.</td></tr><tr><td>DAE</td><td>Food was cold and lacking.</td><td>Excellent service, good food, nice atmosphere and nice atmosphere.</td></tr><tr><td>CU-Gst</td><td>Food was nicely and engaging of flavors.</td><td>Terrible service, not food, and the atmosphere.</td></tr><tr><td>Multi-Div</td><td>Food was prepared, expensive and service.</td><td>Disappointing water, both nice and dry delicious at!</td></tr><tr><td>Style-Emb</td><td>Food was cold and authentic which has.</td><td>Excellent service, great food, and nice atmosphere.</td></tr><tr><td>DeDR</td><td>Food was fresh and great of flavors.</td><td>Terrible service, bad food, and rude atmosphere.</td></tr><tr><td>DAST</td><td>Food was great and clean of flavors.</td><td>Poor service, horrible food, and bland food.</td></tr><tr><td>DAST-C</td><td>Food was delicious and delicious of flavors.</td><td>Terrible service, bad food, and terrible atmosphere.</td></tr><tr><td>PPST</td><td>Food was fresh and exceptional of flavors.</td><td>Horrible service, bad food, and poor service.</td></tr></table> <p>Table 9: TST results on Yelp dataset (sentiment transfer tasks).</p> <table><tr><th>Model</th><th>ACC (%)</th><th>BLEU</th><th>CS</th><th>WD</th><th>F1</th><th>G-Score</th><th>F1-Score</th><th>#Params</th><th>FLOPs</th></tr><tr><td>DeletionOnly</td><td>84.2</td><td>28.7</td><td>0.893</td><td>0.301</td><td>115</td><td>1.80</td><td>0.034</td><td>9.14M</td><td>27.98G</td></tr><tr><td>Template</td><td>78.2</td><td>48.1</td><td>0.550</td><td>0.463</td><td>1959</td><td>1.04</td><td>0.002</td><td>-</td><td>-</td></tr><tr><td>DeletionRetrieval</td><td>88.1</td><td>30.0</td><td>0.897</td><td>0.464</td><td>101</td><td>1.87</td><td>0.039</td><td>9.88M</td><td>29.78G</td></tr><tr><td>B-GST</td><td>89.2</td><td>46.5</td><td>0.939</td><td>0.649</td><td>216</td><td>1.88</td><td>0.018</td><td>-</td><td>-</td></tr><tr><td>G-GST</td><td>72.7</td><td>32.0</td><td>0.867</td><td>0.617</td><td>407</td><td>1.55</td><td>0.019</td><td>-</td><td>-</td></tr><tr><td>PTD</td><td>82.3</td><td>37.4</td><td>0.982</td><td>0.737</td><td>245</td><td>1.94</td><td>0.016</td><td>-</td><td>-</td></tr><tr><td>UTD</td><td>74.0</td><td>41.0</td><td>0.929</td><td>0.448</td><td>394</td><td>1.36</td><td>0.010</td><td>79.97M</td><td>239.64G</td></tr><tr><td>SMILE</td><td>84.4</td><td>14.8</td><td>0.907</td><td>0.294</td><td>210</td><td>1.313</td><td>0.019</td><td>76.27M</td><td>4418.3G</td></tr><tr><td>DRIST</td><td>91.2</td><td>7.6</td><td>0.904</td><td>0.484</td><td>86</td><td>1.41</td><td>0.045</td><td>64.0M</td><td>6.149G</td></tr><tr><td>RET</td><td>83.1</td><td>2.9</td><td>0.827</td><td>0.076</td><td>201</td><td>0.49</td><td>0.013</td><td>-</td><td>-</td></tr><tr><td>CAAE</td><td>82.7</td><td>11.2</td><td>0.901</td><td>0.277</td><td>145</td><td>1.15</td><td>0.027</td><td>12.92M</td><td>0.436G</td></tr><tr><td>ARAE</td><td>83.2</td><td>18.0</td><td>0.874</td><td>0.270</td><td>138</td><td>1.31</td><td>0.028</td><td>4.20M</td><td>5.61G</td></tr><tr><td>CU-Gst</td><td>69.6</td><td>49.5</td><td>0.933</td><td>0.707</td><td>384</td><td>1.69</td><td>0.049</td><td>13.7M</td><td>6.135G</td></tr><tr><td>Multi-Div</td><td>69.6</td><td>17.2</td><td>0.887</td><td>0.244</td><td>299</td><td>0.99</td><td>0.013</td><td>-</td><td>-</td></tr><tr><td>Style-Emb</td><td>47.5</td><td>51.4</td><td>0.926</td><td>0.433</td><td>217</td><td>1.31</td><td>0.018</td><td>-</td><td>-</td></tr><tr><td>DeDR</td><td>79.0</td><td>58.3</td><td>0.97</td><td>0.601</td><td>134</td><td>2.28</td><td>0.030</td><td>44M</td><td>0.862G</td></tr><tr><td>DAST</td><td>90.7</td><td>49.7</td><td>0.961</td><td>0.705</td><td>323</td><td>1.77</td><td>0.012</td><td>26.68M</td><td>0.822G</td></tr><tr><td>DAST-C</td><td>89.6</td><td>41.2</td><td>0.933</td><td>0.540</td><td>450</td><td>1.48</td><td>0.009</td><td>24.31M</td><td>0.508G</td></tr><tr><td>PPST</td><td>85.3</td><td>41.7</td><td>0.902</td><td>0.527</td><td>104</td><td>2.06</td><td>0.038</td><td>34.32M</td><td>76.19G</td></tr></table>	Model	From negative to positive (Yelp)	From positive to negative (Yelp)	Source	Food was cold and lacking of flavors.	Excellent service, and food, and nice atmosphere.	DeletionOnly	They were great, with a wide variety of flavors.	I would give zero stars, when I could.	Template	Favorite of flavors and delicious.	Excellent service, great food, and nice atmosphere.	DeletionRetrieval	Of a large list of flavors.	A waste of time and money.	B-GST	Food was bland and of no flavors.	Great service, great food, and great atmosphere.	G-GST	The food was cold and lacking of flavors.	The service, nice food, and atmosphere.	PTD	Food was fresh and delicious of flavors.	Horrible service, horrible food, and horrible atmosphere.	SMILE		Slow service, good food, and nice atmosphere.	DRIST	Food was delicious and fresh.	Poor service, poor food, and poor atmosphere.	RET	Food is delicious, and the staff.	Service was ok but the food was not hot.	CAAE	Food was fresh and good.	However, food was good, and poor, poor atmosphere.	DAE	Food was cold and lacking.	Excellent service, good food, nice atmosphere and nice atmosphere.	CU-Gst	Food was nicely and engaging of flavors.	Terrible service, not food, and the atmosphere.	Multi-Div	Food was prepared, expensive and service.	Disappointing water, both nice and dry delicious at!	Style-Emb	Food was cold and authentic which has.	Excellent service, great food, and nice atmosphere.	DeDR	Food was fresh and great of flavors.	Terrible service, bad food, and rude atmosphere.	DAST	Food was great and clean of flavors.	Poor service, horrible food, and bland food.	DAST-C	Food was delicious and delicious of flavors.	Terrible service, bad food, and terrible atmosphere.	PPST	Food was fresh and exceptional of flavors.	Horrible service, bad food, and poor service.	Model	ACC (%)	BLEU	CS	WD	F1	G-Score	F1-Score	#Params	FLOPs	DeletionOnly	84.2	28.7	0.893	0.301	115	1.80	0.034	9.14M	27.98G	Template	78.2	48.1	0.550	0.463	1959	1.04	0.002	-	-	DeletionRetrieval	88.1	30.0	0.897	0.464	101	1.87	0.039	9.88M	29.78G	B-GST	89.2	46.5	0.939	0.649	216	1.88	0.018	-	-	G-GST	72.7	32.0	0.867	0.617	407	1.55	0.019	-	-	PTD	82.3	37.4	0.982	0.737	245	1.94	0.016	-	-	UTD	74.0	41.0	0.929	0.448	394	1.36	0.010	79.97M	239.64G	SMILE	84.4	14.8	0.907	0.294	210	1.313	0.019	76.27M	4418.3G	DRIST	91.2	7.6	0.904	0.484	86	1.41	0.045	64.0M	6.149G	RET	83.1	2.9	0.827	0.076	201	0.49	0.013	-	-	CAAE	82.7	11.2	0.901	0.277	145	1.15	0.027	12.92M	0.436G	ARAE	83.2	18.0	0.874	0.270	138	1.31	0.028	4.20M	5.61G	CU-Gst	69.6	49.5	0.933	0.707	384	1.69	0.049	13.7M	6.135G	Multi-Div	69.6	17.2	0.887	0.244	299	0.99	0.013	-	-	Style-Emb	47.5	51.4	0.926	0.433	217	1.31	0.018	-	-	DeDR	79.0	58.3	0.97	0.601	134	2.28	0.030	44M	0.862G	DAST	90.7	49.7	0.961	0.705	323	1.77	0.012	26.68M	0.822G	DAST-C	89.6	41.2	0.933	0.540	450	1.48	0.009	24.31M	0.508G	PPST	85.3	41.7	0.902	0.527	104	2.06	0.038	34.32M	76.19G
Model	From negative to positive (Yelp)	From positive to negative (Yelp)																																																																																																																																																																																																																																																																					
Source	Food was cold and lacking of flavors.	Excellent service, and food, and nice atmosphere.																																																																																																																																																																																																																																																																					
DeletionOnly	They were great, with a wide variety of flavors.	I would give zero stars, when I could.																																																																																																																																																																																																																																																																					
Template	Favorite of flavors and delicious.	Excellent service, great food, and nice atmosphere.																																																																																																																																																																																																																																																																					
DeletionRetrieval	Of a large list of flavors.	A waste of time and money.																																																																																																																																																																																																																																																																					
B-GST	Food was bland and of no flavors.	Great service, great food, and great atmosphere.																																																																																																																																																																																																																																																																					
G-GST	The food was cold and lacking of flavors.	The service, nice food, and atmosphere.																																																																																																																																																																																																																																																																					
PTD	Food was fresh and delicious of flavors.	Horrible service, horrible food, and horrible atmosphere.																																																																																																																																																																																																																																																																					
SMILE		Slow service, good food, and nice atmosphere.																																																																																																																																																																																																																																																																					
DRIST	Food was delicious and fresh.	Poor service, poor food, and poor atmosphere.																																																																																																																																																																																																																																																																					
RET	Food is delicious, and the staff.	Service was ok but the food was not hot.																																																																																																																																																																																																																																																																					
CAAE	Food was fresh and good.	However, food was good, and poor, poor atmosphere.																																																																																																																																																																																																																																																																					
DAE	Food was cold and lacking.	Excellent service, good food, nice atmosphere and nice atmosphere.																																																																																																																																																																																																																																																																					
CU-Gst	Food was nicely and engaging of flavors.	Terrible service, not food, and the atmosphere.																																																																																																																																																																																																																																																																					
Multi-Div	Food was prepared, expensive and service.	Disappointing water, both nice and dry delicious at!																																																																																																																																																																																																																																																																					
Style-Emb	Food was cold and authentic which has.	Excellent service, great food, and nice atmosphere.																																																																																																																																																																																																																																																																					
DeDR	Food was fresh and great of flavors.	Terrible service, bad food, and rude atmosphere.																																																																																																																																																																																																																																																																					
DAST	Food was great and clean of flavors.	Poor service, horrible food, and bland food.																																																																																																																																																																																																																																																																					
DAST-C	Food was delicious and delicious of flavors.	Terrible service, bad food, and terrible atmosphere.																																																																																																																																																																																																																																																																					
PPST	Food was fresh and exceptional of flavors.	Horrible service, bad food, and poor service.																																																																																																																																																																																																																																																																					
Model	ACC (%)	BLEU	CS	WD	F1	G-Score	F1-Score	#Params	FLOPs																																																																																																																																																																																																																																																														
DeletionOnly	84.2	28.7	0.893	0.301	115	1.80	0.034	9.14M	27.98G																																																																																																																																																																																																																																																														
Template	78.2	48.1	0.550	0.463	1959	1.04	0.002	-	-																																																																																																																																																																																																																																																														
DeletionRetrieval	88.1	30.0	0.897	0.464	101	1.87	0.039	9.88M	29.78G																																																																																																																																																																																																																																																														
B-GST	89.2	46.5	0.939	0.649	216	1.88	0.018	-	-																																																																																																																																																																																																																																																														
G-GST	72.7	32.0	0.867	0.617	407	1.55	0.019	-	-																																																																																																																																																																																																																																																														
PTD	82.3	37.4	0.982	0.737	245	1.94	0.016	-	-																																																																																																																																																																																																																																																														
UTD	74.0	41.0	0.929	0.448	394	1.36	0.010	79.97M	239.64G																																																																																																																																																																																																																																																														
SMILE	84.4	14.8	0.907	0.294	210	1.313	0.019	76.27M	4418.3G																																																																																																																																																																																																																																																														
DRIST	91.2	7.6	0.904	0.484	86	1.41	0.045	64.0M	6.149G																																																																																																																																																																																																																																																														
RET	83.1	2.9	0.827	0.076	201	0.49	0.013	-	-																																																																																																																																																																																																																																																														
CAAE	82.7	11.2	0.901	0.277	145	1.15	0.027	12.92M	0.436G																																																																																																																																																																																																																																																														
ARAE	83.2	18.0	0.874	0.270	138	1.31	0.028	4.20M	5.61G																																																																																																																																																																																																																																																														
CU-Gst	69.6	49.5	0.933	0.707	384	1.69	0.049	13.7M	6.135G																																																																																																																																																																																																																																																														
Multi-Div	69.6	17.2	0.887	0.244	299	0.99	0.013	-	-																																																																																																																																																																																																																																																														
Style-Emb	47.5	51.4	0.926	0.433	217	1.31	0.018	-	-																																																																																																																																																																																																																																																														
DeDR	79.0	58.3	0.97	0.601	134	2.28	0.030	44M	0.862G																																																																																																																																																																																																																																																														
DAST	90.7	49.7	0.961	0.705	323	1.77	0.012	26.68M	0.822G																																																																																																																																																																																																																																																														
DAST-C	89.6	41.2	0.933	0.540	450	1.48	0.009	24.31M	0.508G																																																																																																																																																																																																																																																														
PPST	85.3	41.7	0.902	0.527	104	2.06	0.038	34.32M	76.19G																																																																																																																																																																																																																																																														

There is a delicate balance between retaining the original content and removing the original attributes, and existing adversarial models tend to sacrifice one or the other.

It introduces a simple approach to text attribute transfer whose main advantage comes from induction bias, i.e. attributes are usually represented as locally recognisable phrases, which allows it to outperform previous models based on adversarial training. It also reflects some problems in that content and attributes cannot be so cleanly separated by word boundaries.

A fruitful direction would be to develop a concept of attributes that is more general than N-grams, more general than N-grams, but with more inductive bias than arbitrary potential vectors.