

# 1 The evaluation scores for different triple size of Seen dataset

=====

We mean by flat one model from the extended system which contains 3 model:flat,str1,str2  
Where we evaluate just this model in term of different triple size, because the result of other models i.e. str1&str2 for the dataset seen and unseen not very good comparing by this model

We mean by "Modify baseline" the modified baseline of webnlg challeng

## 1.1 Table of Evaluation in term of BLEU

Triple size	Model	Bleu
Triple 1	Modify Baseline	60.73
	Flat	<b>59.26</b>
Triple 2	Modify Baseline	51.77
	Flat	<b>50.47</b>
Triple 3	Modify Baseline	48.60
	Flat	<b>48.32</b>
Triple 4	Modify Baseline	47.29
	Flat	<b>46.97</b>
Triple 5	Modify Baseline	47.24
	Flat	<b>45.15</b>
Triple 6	Modify Baseline	48.01
	Flat	<b>50.75</b>
Triple 7	Modify Baseline	51.88
	Flat	<b>48.95</b>

## 1.2 Table of Evaluation in term of METEOR

Triple size	Model	METEOR
Triple 1	Modify Baseline	0.49
	Flat	<b>0.51</b>
Triple 2	Modify Baseline	0.42
	Flat	<b>0.44</b>
Triple 3	Modify Baseline	0.39
	Flat	<b>0.41</b>
Triple 4	Modify Baseline	0.36
	Flat	<b>0.39</b>
Triple 5	Modify Baseline	0.36
	Flat	<b>0.35</b>
Triple 6	Modify Baseline	0.35
	Flat	<b>0.39</b>
Triple 7	Modify Baseline	0.37
	Flat	<b>0.37</b>

### 1.3 Table of Evaluation in term of TER

Triple size	Model	TER
Triple 1	Modify Baseline	0.26
	Flat	<b>0.25</b>
Triple 2	Modify Baseline	0.39
	Flat	<b>0.37</b>
Triple 3	Modify Baseline	0.45
	Flat	<b>0.41</b>
Triple 4	Modify Baseline	0.48
	Flat	<b>0.47</b>
Triple 5	Modify Baseline	0.48
	Flat	<b>0.50</b>
Triple 6	Modify Baseline	0.46
	Flat	<b>0.43</b>
Triple 7	Modify Baseline	0.42
	Flat	<b>0.44</b>