	From positive to negative (YELP)
SOURCE	my husband got a ruben sandwich , he loved it .
CROSSALIGNED	my husband got a *appetizer* sandwich , *she was it wrong* .
STYLEEMBEDDING	my husband got a *greatest* sandwich , he loved it .
MULTIDECODER	my husband got a *beginning house with however i ignored* .
TEMPLATEBASED	my husband got a ruben sandwich , *don't care*
RETRIEVEONLY	i got the *club* sandwich and *my mom got the chicken salad sandwich* .
DELETEONLY	my husband got a ruben sandwich , *and it was horrible !*
DELETEANDRETRIEVE	my husband got a ruben sandwich , *it was too dry* .
	From factual to humorous (CAPTIONS)
SOURCE	a black and white dog is running through shallow water .
CROSSALIGNED	*two dogs are playing on a field to win the* water .
STYLEEMBEDDING	a black and white dog is running through shallow water .
MULTIDECODER	a black and white dog is running through *grassy* water .
TEMPLATEBASED	a black and white dog is running through shallow water *looking for* .
RETRIEVEONLY	a black and white dog is *slowly* running through *a snowy field* .
DELETEANDRETRIEVE	a black and white dog is running through shallow water *to search for bones* .
DELETEONLY	a black and white dog is running through shallow water *like a fish* .
	From positive to negative (AMAZON)
SOURCE	i would definitely recommend this for a cute case .
CROSSALIGNED	i would *not recommend* this for a *long time* .
STYLEEMBEDDING	i would definitely recommend this for a cute case .
MULTIDECODER	i would definitely recommend this for a *bra does it* .
TEMPLATEBASED	*skip this one* for a cute case .
RETRIEVEONLY	cute *while it lasted . . . so if you want to have a NUM night stand case , this is your* case .
DELETEONLY	i would *not recommend it* for a cute case .
DELETEANDRETRIEVE	i would *not recommend* this for a cute case .

Table 6: Additional example outputs on YELP, CAPTIONS, and AMAZON. Added or changed words are in *italic*. Attribute markers are colored.
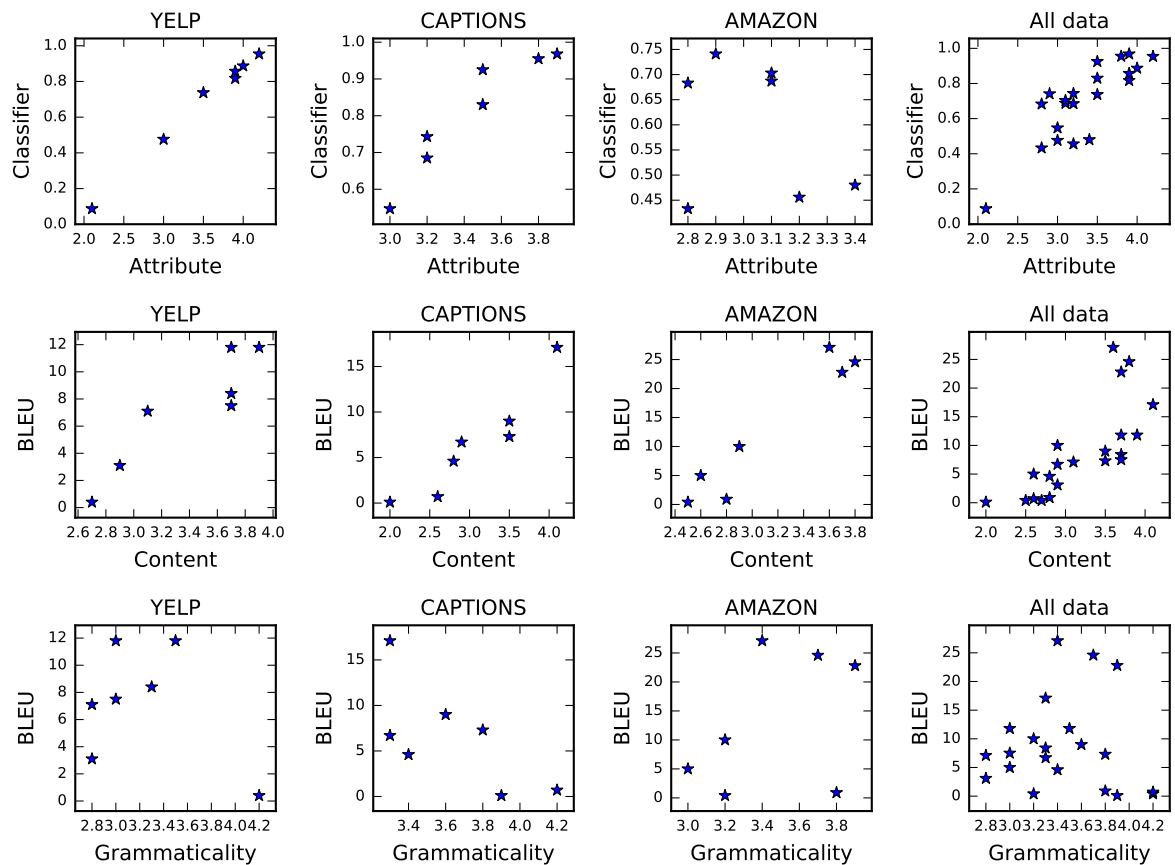
Figure 4: Scatter plots of humans scores vs. automatic metric scores on attribute transfer, content preservation, and grammaticality. The automatic metrics have some correlation with the attribute transfer and content preservation ratings, though results vary across datasets; the metrics do not measure grammaticality.