KVXJ9\_10: nocontra: sc=0, contra: sc=0.2

Nocontra: run stairs somegthing coffee

Contra: be bag walk hold

GT: hold mirror

KVXJ9\_11: nocontra: sc=0, contra: sc=0.166666

Nocontra: awaken stairs run place

Contra: be table walk room

GT: be camera play

This example illustrates contra predicts more accurate than nocontra.

contra:

'G6ZOB\_2618': {'sc': 0.2}: open door open door vs close door

'G6ZOB\_2619': {'sc': 0.0}: open door open door vs turn light

'G6ZOB\_2620': {'sc': 0.0}: open door open door vs turn light

The contra model will predict the same thing in the same video but different qid.

Noncontra:

'G6ZOB\_2618': close door book door

turn light door take

turn light light door

Nocontra

‘8LLI3\_2633': {'sc': 0.5}: close cabinet take door vs close door

'8LLI3\_2634': {'sc': 0.5}: take picture door light vs take picture

'8LLI3\_2635': {'sc': 0.5}: close cabinets read box vs close cabinets

'8LLI3\_2636': {'sc': 0.5}: close door laugh sneeze vs close door

'8LLI3\_2637': {'sc': 0.5}: put shelf sneeze shelf vs put shelf

This example illustrates that in the same video but different qid, the prediction will be affected by the previous video clip.

7IPW7\_3691: nocontra(0.5): walk doorway television read

Contra(0.5): walk doorway get light

GT: walk doorway

7IPW7\_3692: nocontra: walk doorway drink coffee

contra: walk doorway turn light

GT: begin dress

This example shows that contra and nocontra behaves well and the max prediction similarity is 0.5 for both. Nocontra predicts ‘television read’ after ‘walk doorway’ while contra predicts ‘get light’ after ‘walk doorway’. From that being said, two models have different training process and behaves slightly different with the same prediction similarity.

Nocontra:

8LLI3\_2633': {'sc': 0.5}: close cabinet take door vs close door

Nocontra behaves well and the max prediction similarity is 0.5.

Nocontra

'G6ZOB\_2618': {'sc': 0.5}, 'G6ZOB\_2619': {'sc': 0.5}, 'G6ZOB\_2620': {'sc': 0.5}

'OB9D0\_2006': {'sc': 0.5}, 'SUJWY\_168': {'sc': 0.5},

OB9D0\_2007': {'sc': 0.75}: {'pred': 'start camera play read', 'gt': '<sos> start camera play <eos>'}

'SUJWY\_169': {'sc': 0.75}: {'pred': 'begin something laugh eat', 'gt': '<sos> begin something laugh <eos>'},

Contra:

'SUJWY\_169': 0.0 {'pred': 'drink medicine door put', 'gt': '<sos> begin something laugh <eos>'}

'OB9D0\_2007': 0.0 {'pred': 'put picture put shoe', 'gt': '<sos> start camera play <eos>'}

'E89S9\_1513': {'sc': 0.75}, 'E89S9\_1514': {'sc': 0.5},

'E89S9\_1513': {'pred': 'be chair sit eat', 'gt': '<sos> be chair sit <eos>'},

'4CAB5\_1760': {'sc': 0.75}: {'pred': 'be stave cook cook', 'gt': '<sos> be stave cook <eos>'}

'4CAB5\_1761': {'sc': 0.75}: {'pred': 'be stave cook cook', 'gt': '<sos> be stave cook <eos>'}

'7IPW7\_3691': {'pred': 'walk doorway get light', 'gt': '<sos> walk doorway <eos>'}, '7IPW7\_3692': {'pred': 'walk doorway turn light', 'gt': '<sos> walk doorway <eos>'},