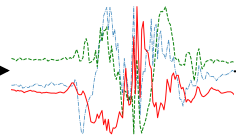
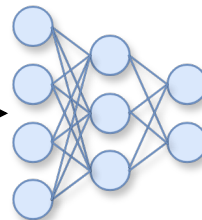


**Wearable Sensor device**



**Sensor Data**

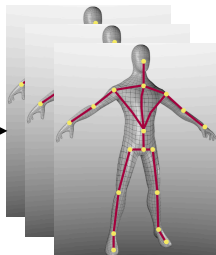


**Classifier Network**

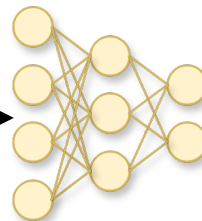
inference →  
"standing",  
"sitting",  
"walking",  
"waving"  
....



**Kinetic Depth Cameras**

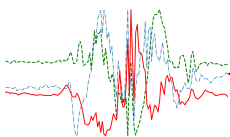


**Skeleton Data**



**Classifier Network**

inference →  
"waving",  
"standing",  
"walking",  
"sitting"  
....

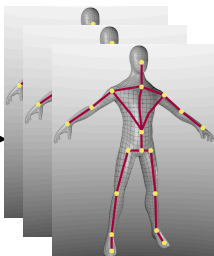


**Sensor Data**

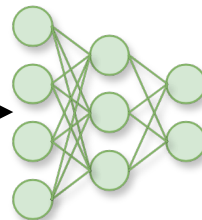
conditioned →

**Skeleton  
Diffusion  
Model**

generate →



**Skeleton Data**



**Classifier Network**

inference →  
"waving",  
"sitting",  
"standing",  
"walking"  
.....