# 3.Implement A Simple Transformer Model

* a) Implement a small transformer with one layer encoder and one layer decoder with self-attention according to the “Attention is All you Need” paper
* b) A sub-layer can be constructed from code available in any package
* c) Show and explain results (input, output) of each sub-layer during training and testingA screenshot of a computer program

  Description automatically generated with low confidence



A screenshot of a computer code

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated with medium confidence

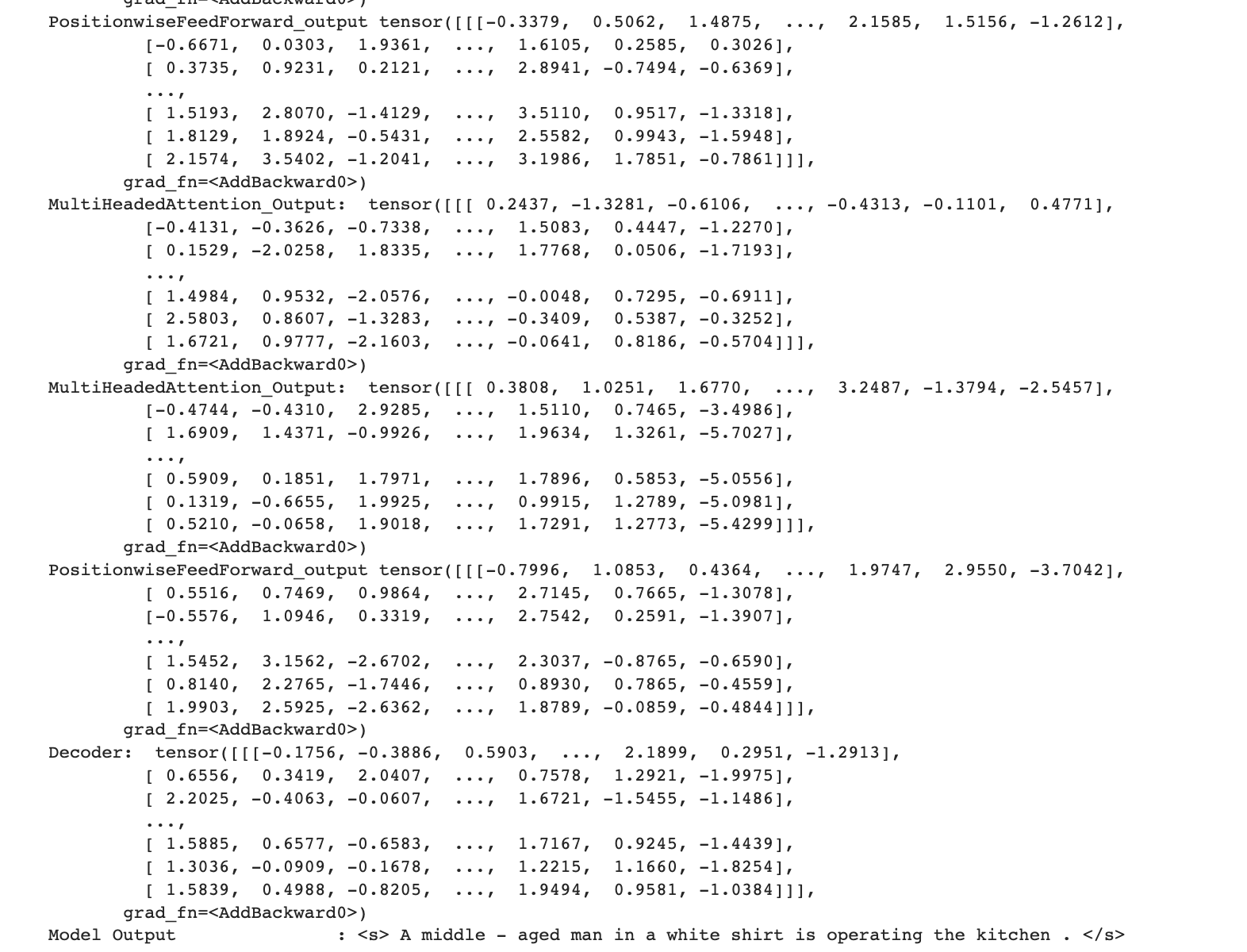
A screenshot of a computer program

Description automatically generated with low confidence

A picture containing text, screenshot, font

Description automatically generated

**Decoder Self Attention**



A picture containing text, screenshot, line

Description automatically generated

A screenshot of a computer code

Description automatically generated with low confidence **Decoder src Attention**

A screenshot of a computer

Description automatically generated with medium confidence

A picture containing pattern, screenshot, line, colorfulness

Description automatically generated