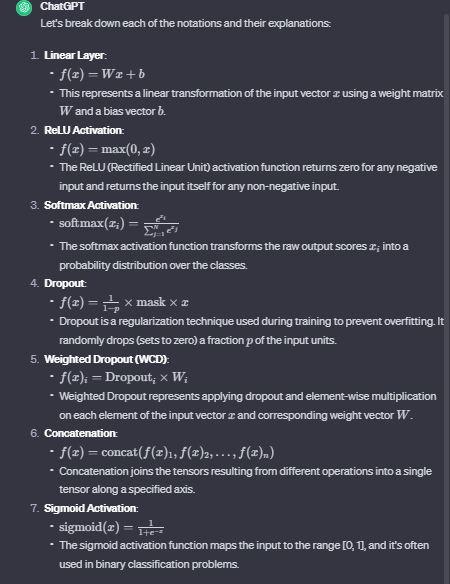
**Rumus**

Notasi

* f(x) is the output vector.
* x is the input vector.
* W is the weight matrix.
* b is the bias vector.
* is the -th element of the input vector x.
* N is the total number of elements in the input vector x.
* e is the base of the natural logarithm (Euler's number), approximately equal to 2.71828.
* is the first linear layer weight for element



Let's break down each of the notations and their explanations:

1. \*\*Linear Layer\*\*:

- \( f(x) = Wx + b \)

- This represents a linear transformation of the input vector \( x \) using a weight matrix \( W \) and a bias vector \( b \).

2. \*\*ReLU Activation\*\*:

- \( f(x) = \max(0, x) \)

- The ReLU (Rectified Linear Unit) activation function returns zero for any negative input and returns the input itself for any non-negative input.

3. \*\*Softmax Activation\*\*:

- \( \text{softmax}(x\_i) = \frac{e^{x\_i}}{\sum\_{j=1}^{N} e^{x\_j}} \)

- The softmax activation function transforms the raw output scores \( x\_i \) into a probability distribution over the classes.

4. \*\*Dropout\*\*:

- \( f(x) = \frac{1}{1-p} \times \text{mask} \times x \)

- Dropout is a regularization technique used during training to prevent overfitting. It randomly drops (sets to zero) a fraction \( p \) of the input units.

5. \*\*Weighted Dropout (WCD)\*\*:

- \( f(x)\_i = \text{Dropout}\_i \times W\_i \)

- Weighted Dropout represents applying dropout and element-wise multiplication on each element of the input vector \( x \) and corresponding weight vector \( W \).

6. \*\*Concatenation\*\*:

- \( f(x) = \text{concat}(f(x)\_1, f(x)\_2, \ldots, f(x)\_n) \)

- Concatenation joins the tensors resulting from different operations into a single tensor along a specified axis.

7. \*\*Sigmoid Activation\*\*:

- \( \text{sigmoid}(x) = \frac{1}{1+e^{-x}} \)

- The sigmoid activation function maps the input to the range [0, 1], and it's often used in binary classification problems.

Notation explanations:

- \( f(x) \) is the output vector.

- \( x \) is the input vector.

- \( W \) is the weight matrix.

- \( b \) is the bias vector.

- \( x\_i \) is the \( i \)-th element of the input vector \( x \).

- \( N \) is the total number of elements in the input vector \( x \).

- \( e \) is the base of the natural logarithm (Euler's number), approximately equal to 2.71828.

- \( W\_i \) is the weight for channel \( i \).