

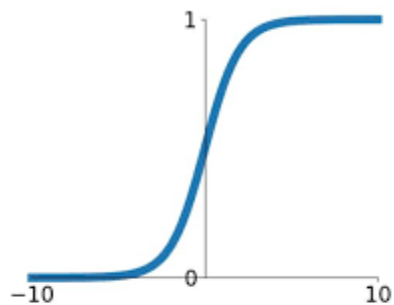
**Activation Function**

# Activation Function

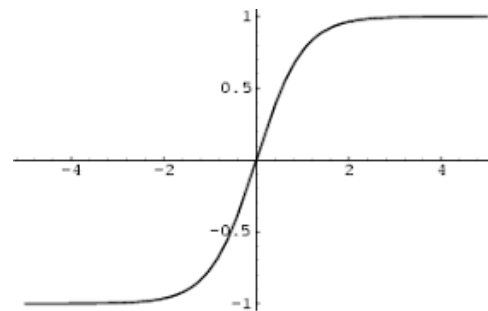
**Activation Function**

# Activation Function

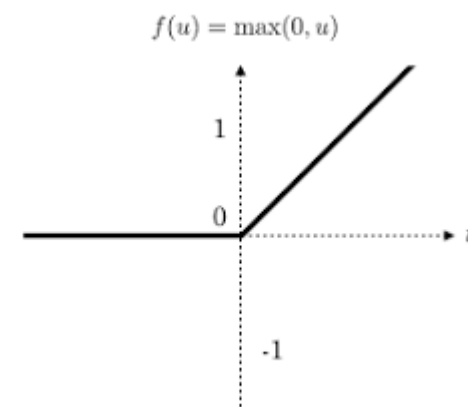
What is the difference?



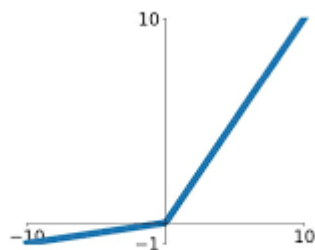
**Sigmoid**



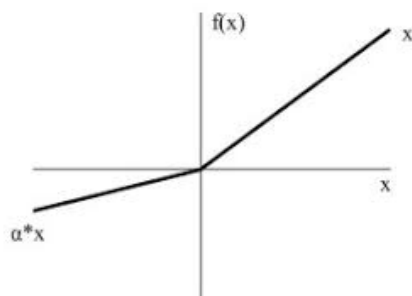
**Hyperbolic Tangent**



**ReLU**

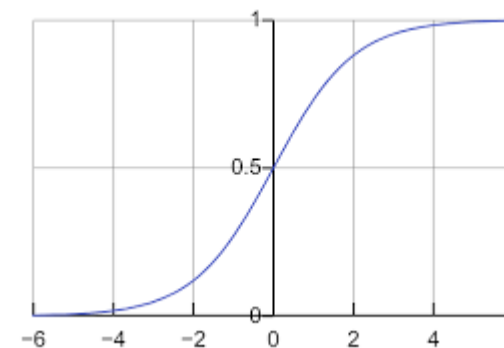
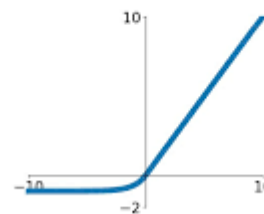


**Leaky ReLU**



**PReLU**

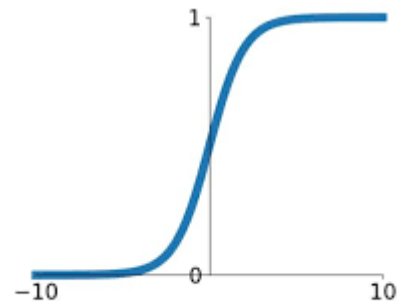
**Exponential Linear Units (ELU)**



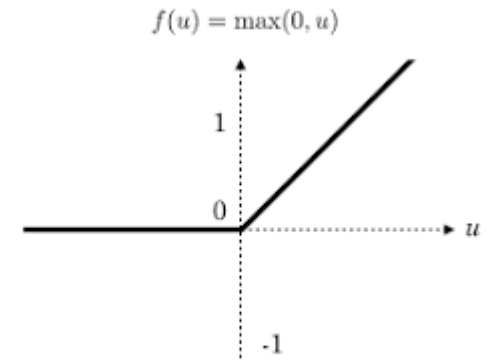
**SoftSign**

# Activation Function

Symmetry



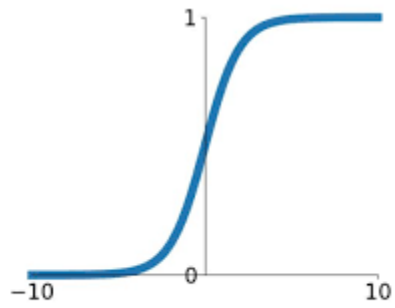
**Sigmoid**



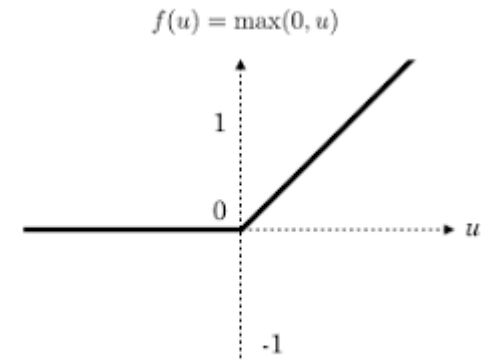
**ReLU**

# Activation Function

Pros and Cons?



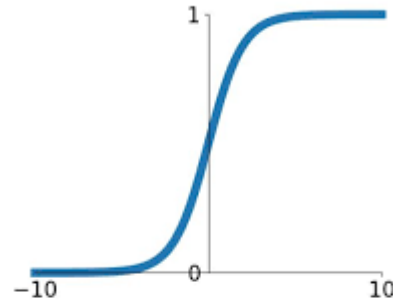
**Sigmoid**



**ReLU**

Pros  
Cons

Sigmoid



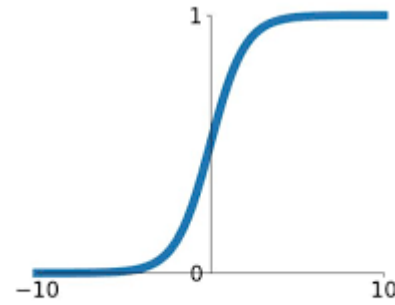
**Sigmoid**

Pros : Symmetry

Cons : Gradient Vanishing

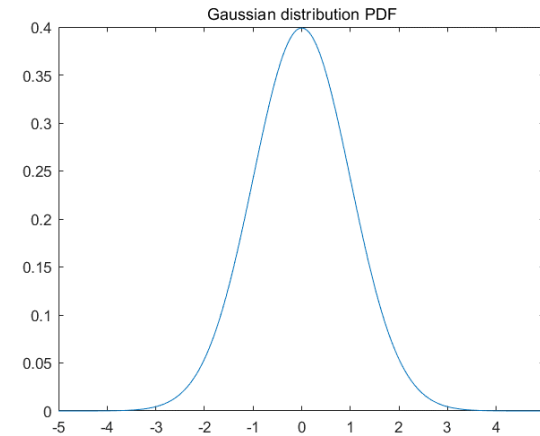
Pros  
Cons

Why is Symmetry an advantage?



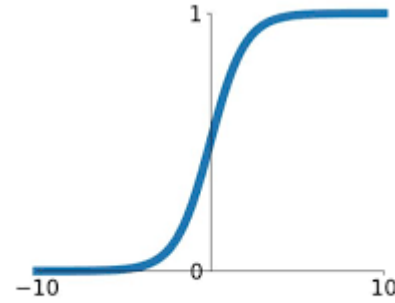
**Sigmoid**

The Gaussian Distribution



Pros  
Cons

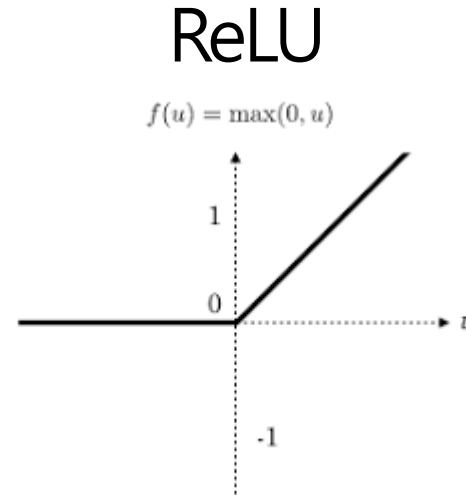
Why is gradient vanishing a disadvantage?



**Sigmoid**



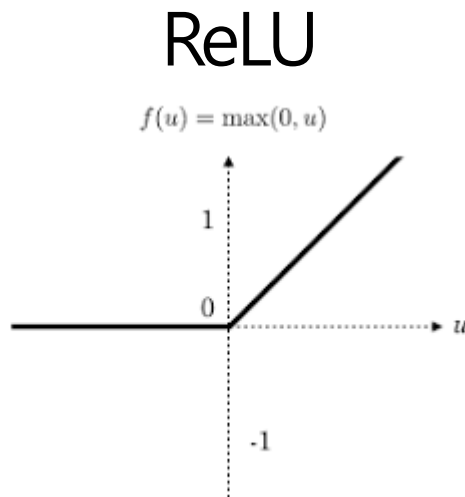
Pros  
Cons



Pros : Fast Caculation, reduce GV

Cons : Baised, Dying ReLU

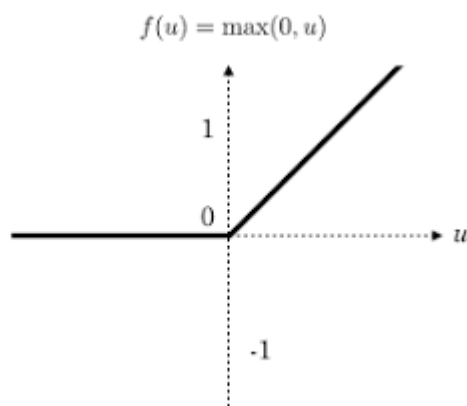
Pros  
Cons



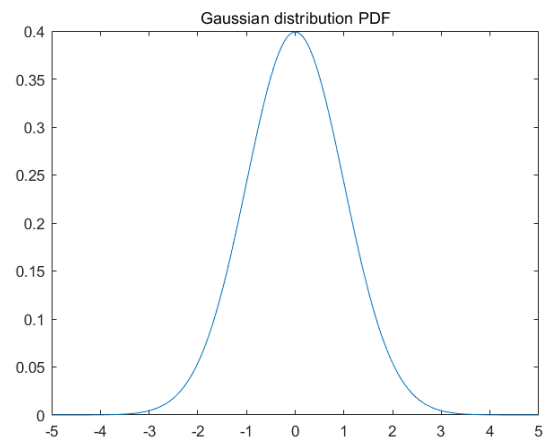
ReLU Gradient :

Pros  
Cons

ReLU



Baised



Pros  
Cons

We can think of two things

First, biased

Second, gradient Vanishing

How do you solve it?

Pros  
Cons

We can think of two things

First, biased

Second, gradient Vanishing

How do you solve it?

감사합니다

THANK YOU