

Result Text Detection

Detected Texts

[deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

-

120%

+

Greedy layer-wise unsupervised pretraining relies on a single-layer auto-encoding learning algorithm such as an RBM, a single-layer auto-encoding model, or another model that learns latent representations. The model is pre-trained using unsupervised learning, taking the output of the pre-trained model and producing as output a new representation of the data, with its relation to other variables, such as categories to predict. See algorithm 15.1 for a formal description.

Greedy layer-wise training procedures based on unsupervised learning have been used to sidestep the difficulty of

[deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

-

120%

+

Greedy
layer-wise
unsupervised
pretraining
relies
on
a
single-layer

representation
learning
algorithm
such
as
an
RBM,
a
single-layer
autoencoder,
a
sparse
coding
model,
or
another
model
that
learns
latent
representations.
Each
layer
is
pretrained
using
unsupervised
learning,
taking
the
output
of

the
previous
layer
and
producing
as
output
a
new
representation
of
the
data,
whose
distribution
(or
its
relation
to
other
variables,
such
as
categories
to
predict)
is
hopefully
simpler.
See
algorithm
15.1
for

a

formal
description.

Greedy

layer-wise

training

procedures

based

on

unsupervised

criteria

have

long

been

used

to

sidestep

the

difficulty

of

[deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

-

120%

+

Greedy layer-wise unsupervised pretraining relies on a single-layer representation learning algorithm such as an RBM, a single-layer autoencoder, or another model that learns latent representations. The model is pretrained using unsupervised learning, taking the input data as input and producing as output a new representation of the data, which can be used to study its relation to other variables, such as categories to predict. See algorithm 15.1 for a formal description.

Greedy layer-wise training procedures based on unsupervised
been used to sidestep the difficulty of
[deeplearningbook.org/contents/rep](http://deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

-

120%

+

Greedy
layer-wise
unsupervised
pretraining
relies
on

a

single-layer
representa-
tion
learning
algorithm
such

as

an

RBM,

a

single-layer
autoencoder,

a

sparse
coding

model,

or

another

model
that
learns
latent
representations.
Each
layer
is
pretrained
using
unsupervised
learning,
taking
the
output
of
the
previous
layer
and
producing
as
output
a
new
representation
of
the
data,
whose
distribution
(or
its

relation
to
other
variables,
such
as
categories
to
predict)
is
hopefully
simpler.
See
algorithm
15.1
for
a
formal
description.
Greedy
layer-wise
training
procedures
based
on
unsupervised
criteria
have
long
been
used
to
sidestep

the

difficulty

of

[deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

-

120%

+

Greedy layer-wise unsupervised pretraining relies on a single-layer representation learning algorithm such as an RBM, a single-layer autoencoder, or another model that learns latent representations. The model is first pretrained using unsupervised learning, taking the output of the first layer as input and producing as output a new representation of the data, which is then used to learn its relation to other variables, such as categories to predict. See algorithm 15.1 for a formal description.

Greedy layer-wise training procedures based on unsupervised learning have been used to sidestep the difficulty of training deep neural networks. See [deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

-

120%

+

Greedy

layer-wise

unsupervised

pretraining

relies

on

a

single-layer

representa-

tation
learning
algorithm
such
as
an
RBM,
a
single-layer
autoencoder,
a
sparse
coding
model,
or
another
model
that
learns
latent
representations.
Each
layer
is
pretrained
using
unsupervised
learning,
taking
the
output
of
the

previous
layer
and
producing
as
output
a
new
representation
of
the
data,
whose
distribution
(or
its
relation
to
other
variables,
such
as
categories
to
predict)
is
hopefully
simpler.
See
algorithm
15.1
for
a

formal
description.

Greedy
layer-wise
training
procedures
based
on
unsupervised
criteria

have
long
been
used
to
sidestep
the
difficulty
of

deeplearningbook.org/contents/rep

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

120%

+

Greedy layer-wise unsupervised pretraining relies on a single
tation learning algorithm such as an RBM, a single-layer auto-
coding model, or another model that learns latent representations
pretrained using unsupervised learning, taking the output of the
and producing as output a new representation of the data, without
its relation to other variables, such as categories to predict. See
See algorithm 15.1 for a formal description.

Greedy layer-wise training procedures based on unsupervised
have been used to sidestep the difficulty of jointly

[deeplearningbook.org/contents/rep](http://deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

120%

+

Greedy

layer-wise

unsupervised

pretraining

relies

on

a

single-layer

represen-

tation

learning

algorithm

such

as

an

RBM,

a

single-layer

autoencoder,

a

sparse

coding

model,

or

another

model

that

learns

latent
representations.
Each
layer
is
pretrained
using
unsupervised
learning,
taking
the
output
of
the
previous
layer
and
producing
as
output
a
new
representation
of
the
data,
whose
distribution
(or
its
relation
to
other

variables,
such
as
categories
to
predict)
is
hopefully
simpler.
See
algorithm
15.1
for
a
formal
description.
Greedy
layer-wise
training
procedures
based
on
unsupervised
criteria
have
long
been
used
to
sidestep
the
difficulty
of

jointly

[deeplearningbook.org/contents/rep](http://deeplearningbook.org/contents/representation-learning)

+

-

120%

+

that greedy layer-wise pretraining is not required to train fully connected architectures, but the unsupervised pretraining approach was successful.

Greedy layer-wise pretraining is called greedy because it is greedy.

526

CHAPTER 15. REPRESENTATION LEARNING

algorithm, meaning that it optimizes each piece of the solution one piece at a time, rather than jointly optimizing all pieces. It is greedy because these independent pieces are the layers of the network.

[deeplearningbook.org/contents/rep](http://deeplearningbook.org/contents/representation-learning)

+

-

120%

+

that

greedy

layer-wise

pretraining

is

not

required

to

train

fully

connected

deep

architectures,
but
the
unsupervised
pretraining
approach
was
the
first
method
to
succeed.

Greedy
layer-wise
pretraining
is
called
greedy
because
it
is
a
greedy
algo-

526

CHAPTER

15.

REPRESENTATION LEARNING

rithm,
meaning
that
it

optimizes
each
piece
of
the
solution
independently,
one
piece
at
a
time,
rather
than
jointly
optimizing
all
pieces.
It
is
called
layer-wise
because
these
independent
pieces
are
the
layers
of
the
network.
Specifically,

greedy

deeplearningbook.org/contents/rep

+

120%

+

Greedy layer-wise pretraining is called greedy because it is

526

CHAPTER 15. REPRESENTATION LEARNING

rithm, meaning that it optimizes each piece of the solution piece at a time, rather than jointly optimizing all pieces. It is greedy because these independent pieces are the layers of the network. Greedy layer-wise pretraining proceeds one layer at a time, training each layer while keeping the previous ones fixed. In particular, the lower layers (bottom layers first) are not adapted after the upper layers are introduced.

Greedy layer-wise pretraining proceeds one layer at a time, training each layer while keeping the previous ones fixed. In particular, the lower layers (bottom layers first) are not adapted after the upper layers are introduced.

deeplearningbook.org/contents/rep

+

120%

+

Greedy

layer-wise

pretraining

is

called

greedy

because

it

is

a

greedy

algo-

526

CHAPTER

15.

REPRESENTATION LEARNING

rithm,
meaning
that
it
optimizes
each
piece
of
the
solution
independently,
one
piece
at
a
time,
rather
than
jointly
optimizing
all
pieces.
It
is
called
layer-wise
because
these
independent
pieces

are
the
layers
of
the
network.
Specifically,
greedy
layer-wise
pretraining
proceeds
one
layer
at
a
time,
training
the
k-th
layer
while
keeping
the
previous
ones
fixed.
In
particular,
the
lower
layers
(which
are

trained
first)
are
not
adapted
after
the
upper
layers
are
introduced.

It
is
called
unsuper-
deeplearningbook.org/contents/rep

+
120%

+

Greedy layer-wise pretraining is called greedy because it is
526

CHAPTER 15. REPRESENTATION LEARNING

rithm, meaning that it optimizes each piece of the solution
piece at a time, rather than jointly optimizing all pieces. It
because these independent pieces are the layers of the netw
greedy

layer-wise pretraining proceeds one layer at a time, training
keeping the previous ones fixed. In particular, the lower lay
first) are not adapted after the upper layers are introduced.
deeplearningbook.org/contents/rep

+
120%

+

Greedy
layer-wise
pretraining
is
called
greedy
because
it
is
a
greedy
algo-

526

CHAPTER

15.

REPRESENTATION LEARNING

rithm,
meaning
that
it
optimizes
each
piece
of
the
solution
independently,
one
piece
at
a
time,

rather
than
jointly
optimizing
all
pieces.
It
is
called
layer-wise
because
these
independent
pieces
are
the
layers
of
the
network.
Specifically,
greedy
layer-wise
pretraining
proceeds
one
layer
at
a
time,
training
the
k-th

layer
while
keeping
the
previous
ones
fixed.

In
particular,
the
lower
layers
(which
are
trained
first)
are
not
adapted
after
the
upper
layers
are
introduced.

It
is
called
unsuper-
deeplearningbook.org/contents/rep
+
120%
+

CHAPTER 15. REPRESENTATION LEARNING

rithm, meaning that it optimizes each piece of the solution piece at a time, rather than jointly optimizing all pieces. It is greedy because these independent pieces are the layers of the network.

layer-wise pretraining proceeds one layer at a time, training each layer while keeping the previous ones fixed. In particular, the lower layers (the first few) are not adapted after the upper layers are introduced. This is motivated because each layer is trained with an unsupervised representation learning algorithm. However, it is not clear how to choose the number of layers.

deeplearningbook.org/contents/representation-learning

+

120%

+

CHAPTER

15.

REPRESENTATION

LEARNING

rithm,

meaning

that

it

optimizes

each

piece

of

the

solution

independently,

one

piece

at

a

time,
rather
than
jointly
optimizing
all
pieces.
It
is
called
layer-wise
because
these
independent
pieces
are
the
layers
of
the
network.
Specifically,
greedy
layer-wise
pretraining
proceeds
one
layer
at
a
time,
training
the

k-th
layer
while
keeping
the
previous
ones
fixed.
In
particular,
the
lower
layers
(which
are
trained
first)
are
not
adapted
after
the
upper
layers
are
introduced.
It
is
called
unsuper-
vised
because
each

layer
is
trained
with
an
unsupervised
representation
learning
algorithm.

However,
it

[deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

120%

+

CHAPTER 15. LEARNING

rithm, meaning that it optimizes each piece of the solution
piece at a time, rather than jointly optimizing all pieces. It
because these independent pieces are the layers of the netw
greedy

layer-wise pretraining proceeds one layer at a time, training
keeping the previous ones fixed. In particular, the lower lay
first) are not adapted after the upper layers are introduced.
vised because each layer is trained with an unsupervised re
algorithm. However, it

[deeplearningbook.org/contents/rep](https://www.deeplearningbook.org/contents/representation.html)

+

[https://www.deeplearningbook.org/contents/representation.h](https://www.deeplearningbook.org/contents/representation.html)

120%

+

CHAPTER

15.

LEARNING

rithm,
meaning
that
it
optimizes
each
piece
of
the
solution
independently,
one
piece
at
a
time,
rather
than
jointly
optimizing
all
pieces.
It
is
called
layer-wise
because
these
independent
pieces
are
the

layers
of
the
network.
Specifically,
greedy
layer-wise
pretraining
proceeds
one
layer
at
a
time,
training
the
k-th
layer
while
keeping
the
previous
ones
fixed.
In
particular,
the
lower
layers
(which
are
trained
first)

are
not
adapted
after
the
upper
layers
are
introduced.
It
is
called
unsuper-
vised
because
each
layer
is
trained