Package 'ProteinDescriptors'

March 3, 2016

Type Package

Title Generates Various Protein Descriptors for Machine Learning Algorithms

Version 0.1.0

Description An implementation of protein descriptors in R. These descriptors combine the advantages of being fixed length and including partial sequential effect: Various length of protein sequences are described with fixed length vectors that are suitable for machine learning algorithms, and still includes partial sequential effect.

License BSD 3-clause License + file LICENSE

LazyData TRUE

RoxygenNote 5.0.1

Suggests testthat

NeedsCompilation no

Author Berat Kurar [aut, cre],
Sami Abu Sneineh [aut],
Yaqoub Ashhab [aut]

Maintainer Berat Kurar <beratkurar@gmail.com>

Repository CRAN

Date/Publication 2016-03-03 23:40:47

R topics documented:

`DiscreteBlosum` *Discrete blosum descriptor.*

Description

`DiscreteBlosum` returns the sum of blosum descriptors of amino acids in a protein sequence.

Usage

```
DiscreteBlosum(x)
```

Arguments

`x`	A string of amino acid letters

Value

A 20 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteBlosum(x)
```

`DiscreteSequentialBlosumTwoParts` *Discrete sequential blosum descriptor with split number=2.*

Description

`DiscreteSequentialBlosumTwoParts` returns the concatenation of the sum of blosum descriptors of amino acids in each split of a protein sequence.

Usage

```
DiscreteSequentialBlosumTwoParts(x)
```

Arguments

`x`	A string of amino acid letters

Value

A 40 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSequentialBlosumTwoParts(x)
```

`DiscreteSequentialSparseFiveParts` *Discrete sequential sparse descriptor with split number=5.*

Description

`DiscreteSequentialSparseFiveParts` returns the concatenation of the sum of sparse descriptors of amino acids in each split of a protein sequence.

Usage

```
DiscreteSequentialSparseFiveParts(x)
```

Arguments

`x`	A string of amino acid letters

Value

A 100 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSequentialSparseFiveParts(x)
```

DiscreteSequentialSparseFourParts

Discrete sequential sparse descriptor with split number=4.

Description

DiscreteSequentialSparseFourParts returns the concatenation of the sum of sparse descriptors of amino acids in each split of a protein sequence.

Usage

```
DiscreteSequentialSparseFourParts(x)
```

Arguments

x	A string of amino acid letters

Value

A 80 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSequentialSparseFourParts(x)
```

DiscreteSequentialSparseThreeParts

Discrete sequential sparse descriptor with split number=3.

Description

DiscreteSequentialSparseThreeParts returns the concatenation of the sum of sparse descriptors of amino acids in each split of a protein sequence.

Usage

```
DiscreteSequentialSparseThreeParts(x)
```

Arguments

x	A string of amino acid letters

Value

A 60 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSequentialSparseThreeParts(x)
```

`DiscreteSequentialSparseTwoParts`
Discrete sequential sparse descriptor with split number=2.

Description

`DiscreteSequentialSparseTwoParts` returns the concatenation of the sum of sparse descriptors of amino acids in each split of a protein sequence.

Usage

```
DiscreteSequentialSparseTwoParts(x)
```

Arguments

`x`	A string of amino acid letters

Value

A 40 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSequentialSparseTwoParts(x)
```

`DiscreteSequentialTdTwoParts`
Discrete sequential 3D descriptor with split number=2.

Description

`DiscreteSequentialTdTwoParts` returns the concatenation of the sum of 3D descriptors of amino acids in each split of a protein sequence.

Usage

```
DiscreteSequentialTdTwoParts(x)
```

Arguments

x	A string of amino acid letters

Value

A 6 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSequentialTdTwoParts(x)
```

DiscreteSparse *Discrete sparse descriptor.*

Description

DiscreteSparse returns the sum of sparse descriptors of amino acids in a protein sequence.

Usage

```
DiscreteSparse(x)
```

Arguments

x	A string of amino acid letters

Value

A 20 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteSparse(x)
```

DiscreteTd *Discrete 3D descriptor.*

Description

DiscreteTd returns the sum of 3D descriptors of amino acids in a protein sequence.

Usage

```
DiscreteTd(x)
```

Arguments

x	A string of amino acid letters

Value

A 3 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
DiscreteTd(x)
```

IncrementalDiscreteSparse *Incremental discrete sparse descriptor.*

Description

incrementalDiscreteSparse returns the sum of incremented sparse descriptors of amino acids in a protein sequence.

Usage

```
IncrementalDiscreteSparse(x)
```

Arguments

x	A string of amino acid letters

Value

A 20 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
IncrementalDiscreteSparse(x)
```

LastFirstDiscreteSparse *Last first discrete sparse descriptor.*

Description

LastFirstDiscreteSparse returns the concatenation of the sum of sparse descriptors of amino acids in a protein sequence and the sum of combination of last part of an amino acid descriptor with the first part of its neighbour amino acid descriptor.

Usage

```
LastFirstDiscreteSparse(x)
```

Arguments

x	A string of amino acid letters

Value

A 40 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
LastFirstDiscreteSparse(x)
```

SequentialBlosum *Sequential blosum descriptor.*

Description

SequentialBlosum returns the concatenation of blosum descriptors of amino acids in a protein sequence.

Usage

```
SequentialBlosum(x)
```

Arguments

x	A string of amino acid letters

Value

A 20*n dimensional numeric vector where n is the protein length

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
SequentialBlosum(x)
```

SequentialDiscreteBlosum

Sequential discrete blosum descriptor with step size=2.

Description

SequentialDiscreteBlosum returns the sum of the concatenation of blosum descriptors of amino acids at every step size in a protein sequence.

Usage

```
SequentialDiscreteBlosum(x)
```

Arguments

x	A string of amino acid letters

Value

A 40 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
SequentialDiscreteBlosum(x)
```

SequentialDiscreteSparse

Sequential discrete sparse descriptor with step size=2.

Description

SequentialDiscreteSparse returns the sum of the concatenation of sparse descriptors of amino acids at every step size in a protein sequence.

Usage

```
SequentialDiscreteSparse(x)
```

Arguments

x	A string of amino acid letters

Value

A 40 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
SequentialDiscreteSparse(x)
```

SequentialDiscreteTd *Sequential discrete 3D descriptor with step size=2.*

Description

SequentialDiscreteTd returns the sum of the concatenation of 3D descriptors of amino acids at every step size in a protein sequence.

Usage

```
SequentialDiscreteTd(x)
```

Arguments

x	A string of amino acid letters

Value

A 6 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
SequentialDiscreteTd(x)
```

SequentialSparse *Sequential sparse descriptor.*

Description

SequentialSparse returns the concatenation of sparse descriptors of amino acids in a protein sequence.

Usage

```
SequentialSparse(x)
```

Arguments

x	A string of amino acid letters

Value

A 20*n dimensional numeric vector where n is the protein length

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
SequentialSparse(x)
```

SequentialTd *Sequential 3D descriptor.*

Description

SequentialTd returns the concatenation of 3D descriptors of amino acids in a protein sequence.

Usage

```
SequentialTd(x)
```

Arguments

x	A string of amino acid letters

Value

A 3*n dimensional numeric vector where n is the protein length

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
SequentialTd(x)
```

ShiftedDiscreteSparse *Discrete sparse descriptor.*

Description

DiscreteSparse returns the sum of sparse descriptors of amino acids in a protein sequence.

Usage

```
ShiftedDiscreteSparse(x)
```

Arguments

x	A string

Value

A 20 dimensional numeric vector

Examples

```
x = "LALHLLLLHMHMMDRSLLLH"
ShiftedDiscreteSparse(x)
```

Index