

# 患者状態表現の 病名交換コードへのマッピング

柴田大作<sup>1</sup>, 嶋本公德<sup>1</sup>, 篠原恵美子<sup>1</sup>, 河添悦昌<sup>1</sup>

1: 東京大学大学院 医学系研究科 医療AI開発学講座

# 背景

- テキストから抽出した患者の状態表現を医学用語集へマッピングすることは重要
  - 文字列一致のみでは十分な精度が得られない可能性
  - 機械学習を用いる場合にもいくつかの問題点
    - アノテーションされた学習データが必要
    - 状態表現の全てがマッピングされるとは限らない

主訴: **嘔気** 現病歴: スポーツ好きで、毎年スキーに行っていた。約5年前に**スキーで転倒しやすくなった**ことに**気づき**、精神内科を受診したところ、**SBMA**と診断された。

**眼球運動**は**全方向で追従**が**できず**、頭位眼球運動反射は保たれていた。**嚔下**が**悪く**、**流涎**が顕著。

現病歴: **重い物の持ち上げ**が**困難**になり、**階段の昇り**が**遅くなる**など**四肢**の**筋力低下**が進行した。



状態表現	標準病名	コード (約2万クラス)
嘔気	嘔気	UAVC
スキー/転倒しやすくなった/気づき	-	x
SBMA	球脊髄性筋萎縮症	DCPR
眼球運動/全方向で追従/できず	眼球運動障害	N3SU
嚔下/悪く	嚔下障害	RHD7
流涎	流涎症	FHC4
重い物の持ち上げ/困難	-	x
階段の昇り/遅くなる	-	x
四肢/筋力低下	四肢筋力低下	HPJL



- 症状や所見などの頻度集計
- 特定の疾患を有する症例の検索



固有表現と固有表現間の関係の情報から  
ルールベースによる患者状態表現の抽出

用語集 (標準病名マスター)へのマッピング

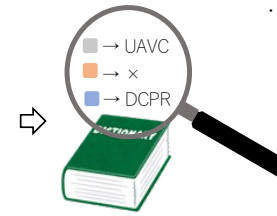
結果の利活用

・症状や所見などの頻度集計  
 ・特定の疾患を有する症例の検索 ..etc

# 目的

主訴: 嘔気  
 現病歴: スポーツ好きで、  
 毎年スキーに行っていた。  
 約5年前にスキーで転倒  
 しやすくなったことに気  
 づき、精神内科を受診し  
 たところ、**球脊髄性筋萎  
 縮症**と診断された。

診療テキストからの  
 患者状態表現の抽出



用語集へのマッピング



結果の利活用

- 機械学習による患者の状態表現の病名交換コード（コード）へのマッピング
  - 学習データには既存の医学用語集のみを使用
    - 学習データの作成コストが大きい問題に対応
  - k最近傍法（近傍法）を用いた病名交換コードへのマッピング
    - 用語集に収載されていない表現にも対応

# 方法: 学習データの作成

標準病名マスター<sup>1</sup> (100,536件, 26,054件のユニークなコード)  
 病名基本テーブルを索引テーブルで拡張  
 1コードあたり平均で3.9単語

病名表記	病名交換コード
1型糖尿病	T48P

病名表記と標準病名が  
一致するレコードを検索

万病辞書<sup>2</sup> (47,930件)  
 信頼度レベルがS~Aの出現形のみを使用

出現形	標準病名	信頼度レベル
急性発症 1 型糖尿病	1型糖尿病	A
特発性 1 型糖尿病	1型糖尿病	A
I D D M	1型糖尿病	A
自己免疫性 1 型糖尿病	1型糖尿病	A

出現形を同義語として  
追加

学習データ (122,492件)  
 1コードあたり平均で4.7単語

病名表記	病名交換コード
1型糖尿病	T48P
急性発症 1 型糖尿病	T48P
特発性 1 型糖尿病	T48P
I D D M	T48P
自己免疫性 1 型糖尿病	T48P

1: ICD10対応標準病名マスター: <https://www2.medis.or.jp/stdcd/byomei/index.html>

2: Kaoru Ito, Hiroyuki Nagai, Taro Okahisa, Shoko Wakamiya, Tomohide Iwao, Eiji Aramaki: J-MeDic: A Japanese Disease Name Dictionary based on Real Clinical Usage, International Conference on Language Resources and Evaluation (LREC) (Japan, Miyazaki), pp. 2365-2369, 2018.

# 方法: 評価データの作成

状態表現	標準病名	コード	対象
SBMA	球脊髄性筋萎縮症	DCPR	✓
ろれつが回りにくく	言語障害	VEA6	✓
発汗 なかった	-	UNK	✓
肺疾患 存在なく	-	UNK	✓
血圧 100mmHg	-	UNK	✓
嘔気	嘔気	UAVC	✓

- 症例報告コーパス<sup>3</sup>に出現する状態表現に対して人手でコードを付与
  - 対応するコードがない場合は医学用語集に収載されていないことを示すUNKを付与
  - **機械学習モデルの評価にのみ使用**



16,520件の状態表現に  
アノテーション

- コードが付与: 6,860件
- UNKが付与: 9,660件



状態表現	標準病名	コード
ろれつが回りにくく	言語障害	VEA6
発汗 なかった	-	UNK
肺疾患 存在なく	-	UNK
血圧 100mmHg	-	UNK
嘔気	嘔気	UAVC

- 状態表現と標準病名が完全一致する表現を削除 (3,568件)
- 重複を削除 (2,285件)

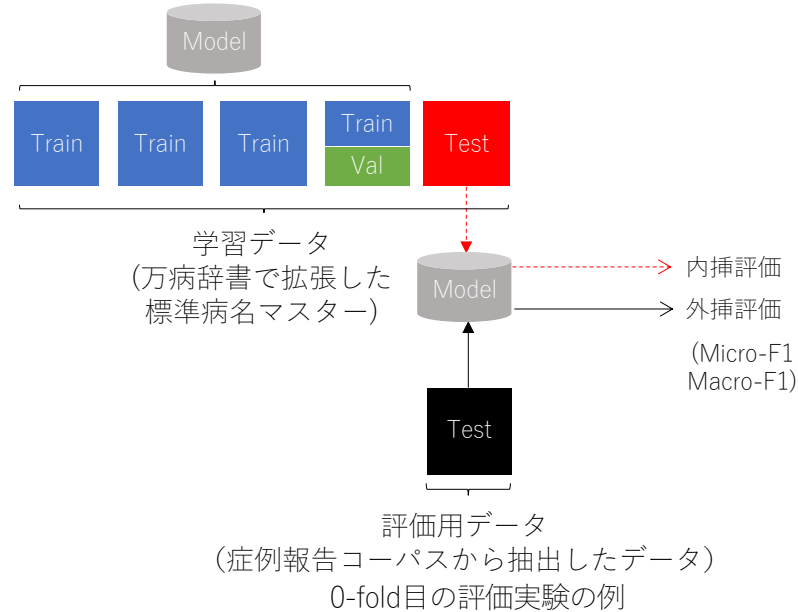


状態表現	標準病名	コード
ろれつが回りにくく	言語障害	VEA6
発汗 なかった	-	UNK
肺疾患 存在なく	-	UNK
血圧 100mmHg	-	UNK
嘔気	嘔気	UAVC

最終的な評価用データは**10,667**件

- コードが付与: 2,208件
- UNKが付与: 8,459件

# 方法: 評価実験



## • 5分割交差検証による評価

### • 内挿評価

- 学習データのみを使用して評価 (122,492件)

### • 外挿評価

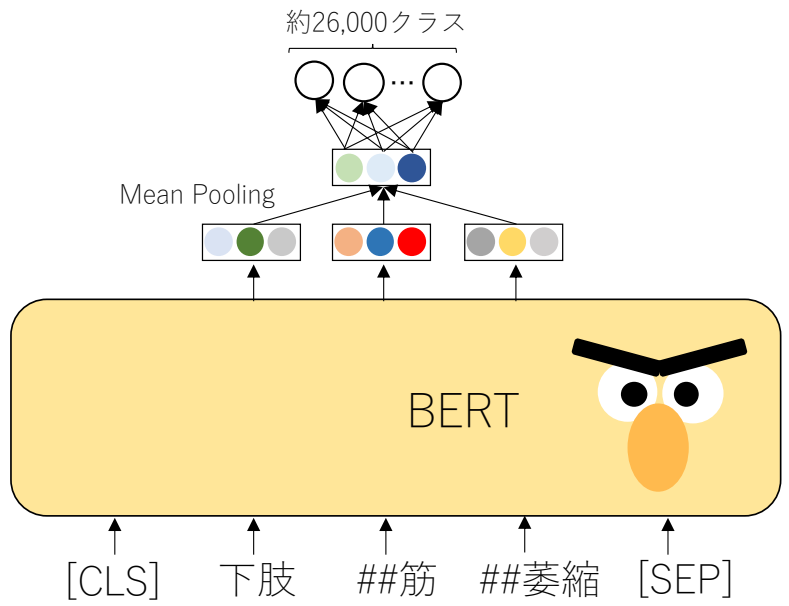
- 学習データで学習したモデルを用いて評価データで評価
  - UNKを含めない外挿評価 (2,208件)
    - 評価データからコードがUNKのデータを削除
  - UNKを含める外挿評価 (10,667件)
    - 全ての評価データを使用

## 方法: 機械学習モデル

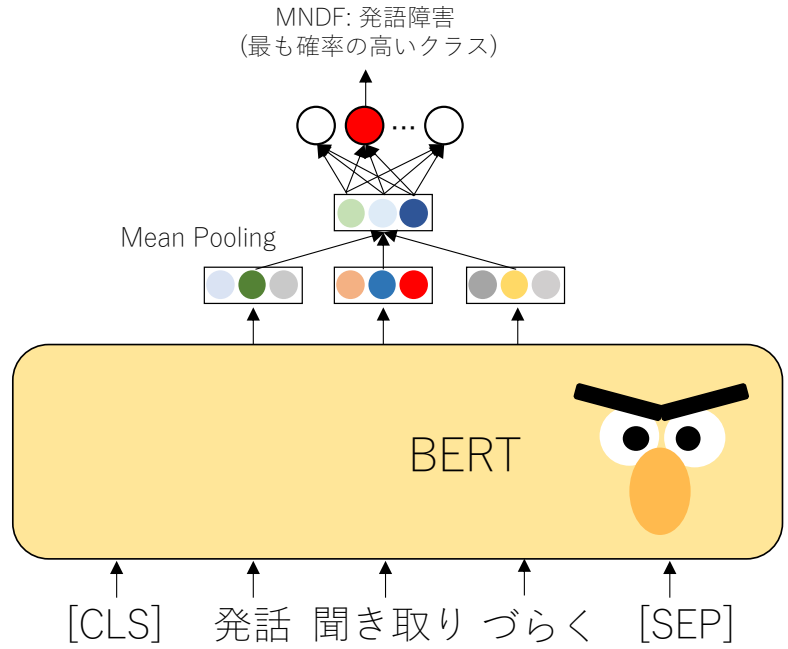
- 診療テキストで事前学習されたUTH-BERT<sup>4</sup>をベースとした2つの手法を使用
  - 分類法
    - BERTの最終層に線形層を接続し, 適切なコードを予測
  - 近傍法
    - 分類法で学習した分類モデルを用いて訓練データの埋め込み表現を取得し, k最近傍法を学習
    - テストデータでも同様に埋め込み表現を取得し, k=1とk=3の場合でコードを予測

# 方法: 分類法

## ① 分類モデルを学習



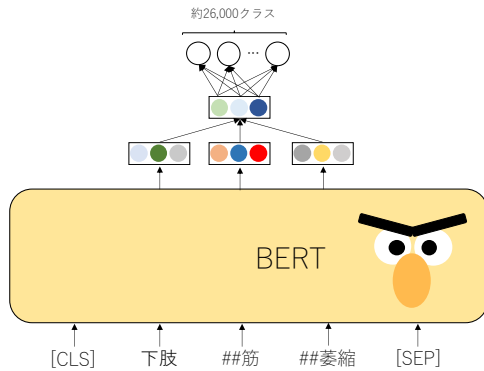
## ② テストデータを予測



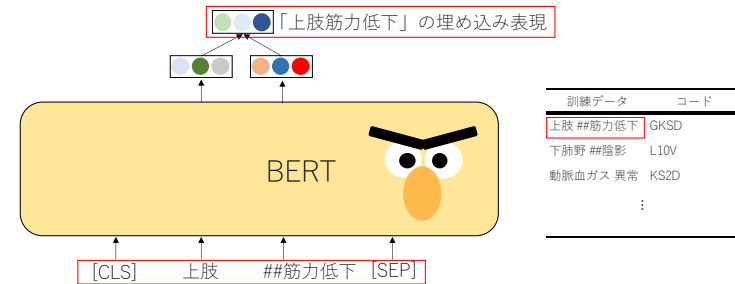


# 方法: UNKを含めない近傍法

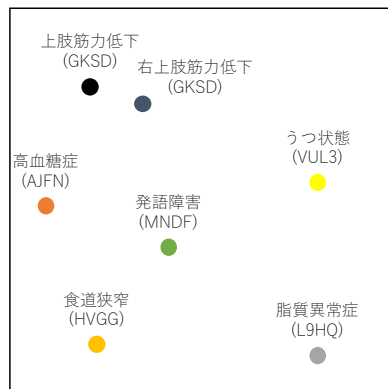
## ① 分類モデルを学習



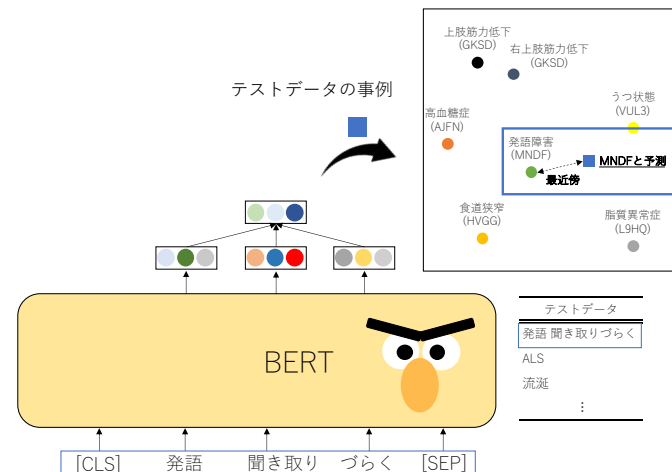
## ② 訓練データの埋め込み表現を得る



## ③ 近傍法を学習

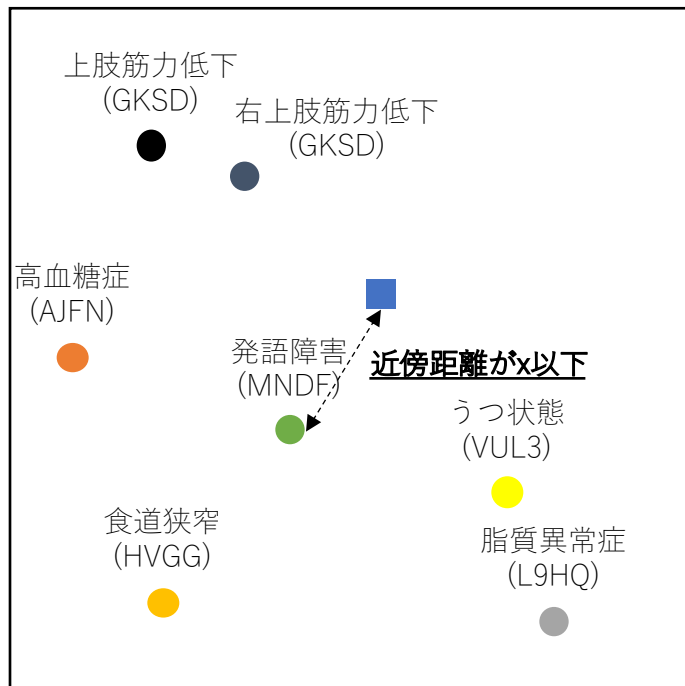


## ④ テストデータを予測

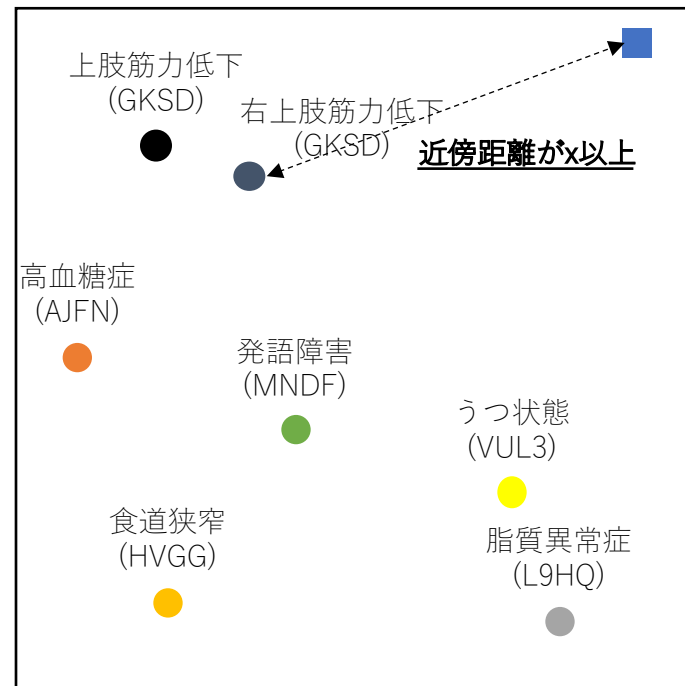


# 方法: UNKを含める場合の近傍法

■ はテストデータの事例



最近傍との距離がx以下  
ならば最近傍のコードを付与



最近傍との距離がx以上  
ならばUNKを付与

※ モデルはUNKを含めない外挿評価で最も精度が高かったモデルを使用

## 方法: 訓練データの拡張

- ティ辞書<sup>5</sup>を用いて訓練データを拡張
  - 優先語, 上位語, 下位語, 同義語などの階層構造を有した辞書
  - 事例が辞書に優先語としてあればその同義語を, 同義語としてあれば他の同義語を訓練データに追加
  - 平均で訓練データが21,696単語増加

項目	用語
優先語	心疾患
同義語	心臓疾患
	心障害
	心臓障害
	心臓病

※ ティ辞書の例

訓練データ: 心疾患

拡張した訓練データ: 心疾患, 心臓疾患, 心障害, 心臓障害, 心臓病

訓練データ: 心臓疾患

拡張した訓練データ: 心臓疾患, 心障害, 心臓障害, 心臓病

# 結果: 内挿評価とUNKを含めない外挿評価

種類	内挿評価								UNKを含めない外挿評価							
	SimString		分類法		近傍法				SimString		分類法		近傍法			
データ拡張	×	○	×	○	×	○	×	○	×	○	×	○	×	○	×	○
近傍数	-	-	-	-	k=1	k=3	k=1	k=3	-	-	-	-	k=1	k=3	k=1	k=3
Micro-F1	0.296	0.295	0.616	0.629	0.633	0.560	<b>0.649</b>	0.574	0.153	0.153	0.499	0.522	0.521	0.433	<b>0.554</b>	0.469
Macro-F1	0.210	0.210	0.438	0.451	0.462	0.381	<b>0.478</b>	0.394	0.174	0.174	0.353	0.375	0.371	0.281	<b>0.403</b>	0.310

- 内挿評価
  - 1近傍法でデータ拡張を行った場合の精度が最も高い
- UNKを含めない外挿評価
  - 1近傍法でデータ拡張を行った場合の精度が最も高い

## 結果: UNKを含める外挿評価

近傍距離	評価指標の計算にUNKを含める			評価指標の計算にUNKを含めない	
	Micro-F1	Macro-F1	UNKのF1	Micro-F1	Macro-F1
2	0.840	0.258	0.910	0.351	0.258
4	0.849	0.307	0.916	0.409	0.306
6	0.864	0.386	0.925	0.494	0.385
8	<b>0.871</b>	0.467	<b>0.933</b>	<b>0.556</b>	<b>0.466</b>
10	0.797	<b>0.499</b>	0.875	0.518	0.498
12	0.403	0.461	0.407	0.394	0.461

※ モデルは1近傍法を使用  
最近傍の事例までの距離が近傍距離を超える場合は予測をUNKとする

- 評価指標の計算にUNKを含める場合
  - $x=8$ の時にMicro-F1が、 $x=10$ の時にMacro-F1が最大
- 評価指標の計算にUNKを含めない場合
  - $x=8$ の時にMicro-F1が、Macro-F1共に最大

# 結果: 事例分析

## 1近傍法の結果

UNKを含めない外挿評価の事例分析								
状態表現	標準病名マスター				ICD-10		一致する桁数	数
	正解		予測		正解	予測		
	コード	病名表記	コード	病名表記	ICD-10			
下肢力入らなく	SRF1	下肢脱力感	NU8T	筋脱力	R298	R298	全桁	174
水泡音	SPA1	湿性ラ音	FDH7	ラ音	R098	R098		
手指伸展	U3PL	伸展拘縮	M1T4	手指伸展拘縮	M2459	M2454	4桁目	4
下肢疼痛	ES36	下腿痛	BBPM	下肢痛	M7966	M7969		
IP	SSA5	間質性肺炎	K5A7	通常型間質性肺炎	J849	J841	3桁目	133
SMA II 型	G9QP	脊髄性筋萎縮症 I I 型	KLR1	脊髄性筋萎縮症 I 型	G121	G120		
ふらつき	VNHQ	運動失調	JV2N	歩行異常	R270	R268	2桁目	74
NAFLD	NGK1	非アルコール性脂肪性肝疾患	DL9H	非アルコール性脂肪性肝炎	K760	K758		
小脳皮質萎縮症	NPUB	皮質性小脳萎縮症	V9BN	亜急性小脳変性症	G112	G319	1桁目	115
精神症状	DTS7	精神障害	AQ3A	精神病	F99	F29		
ろれつが回りにくく むせる	VEA6 NDGS	言語障害 異物誤嚥	AUKS AD6D	構音障害 嚥下困難	F809 T179	R471 R13	一致しない	482

## 考察

種類	内挿評価								UNKを含めない外挿評価							
	SimString		分類法		近傍法				SimString		分類法		近傍法			
データ拡張	×	○	×	○	×		○		×	○	×	○	×		○	
近傍数	-	-	-	-	k=1	k=3	k=1	k=3	-	-	-	-	k=1	k=3	k=1	k=3
Micro-F1	0.296	0.295	0.616	0.629	0.633	0.560	<b>0.649</b>	0.574	0.153	0.153	0.499	0.522	0.521	0.433	<b>0.554</b>	0.469
Macro-F1	0.210	0.210	0.438	0.451	0.462	0.381	<b>0.478</b>	0.394	0.174	0.174	0.353	0.375	0.371	0.281	<b>0.403</b>	0.310

### • 分類法と近傍法

- 同じ実験条件下では近傍法の方が精度が高い
- 1コードあたり学習データが平均で4.7語と少ない
  - BERTのfine-tuningはできているが、十分に分類法(線形層)のパラメーターを学習できなかった可能性
  - BERTのみを活用する近傍法が有効だった可能性

### • データ拡張の効果

- 分類法と近傍法の両方で精度が向上
- 学習データの増加による更なる精度改善が期待

# まとめと今後の予定

- まとめ
  - 標準病名マスターを用いた分類法と近傍法による状態表現の病名交換用コードへのマッピング
  - マッピングにおける近傍法の有効性を確認
- 今後の予定
  - 距離学習などを用いたより良い埋め込み表現の獲得
  - スパンベースのモデルによる評価