FORMULA: FactORized MUlti-task LeArning for task discovery in personalized medical models

Jianpeng Xu[+], Jiayu Zhou[§], Pang-Ning Tan[+]

[+] Department of Computer Science & Engineering, Michigan State University
[§] Samsung Research America

Outline

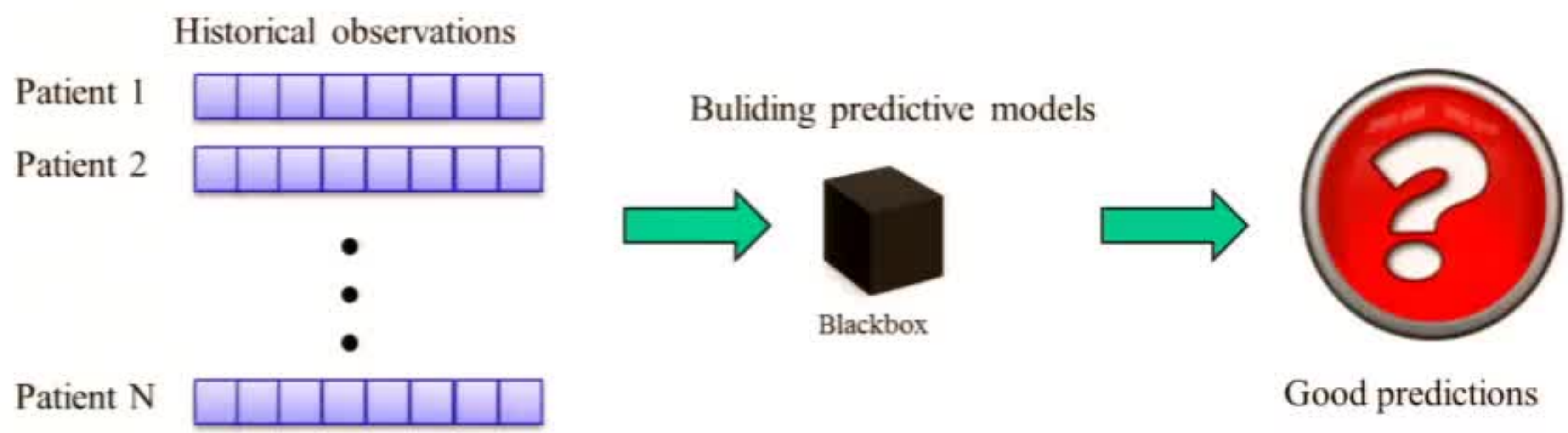
MICHIGAN

- Personalized medical predictive modeling
- Related works
- The proposed FORMULA Framework
- Experiments and Results
- Conclusion

Personalized Medical Predictive Modeling

MICHIGAN

- The amount of patient record data is growing with the development and availability of electronic medical records (EMR)
- The patients' medical records are leveraged by practitioners in applications for clinical decision support and care management systems by performing various predictive modeling tasks for risk predictions and disease analysis.

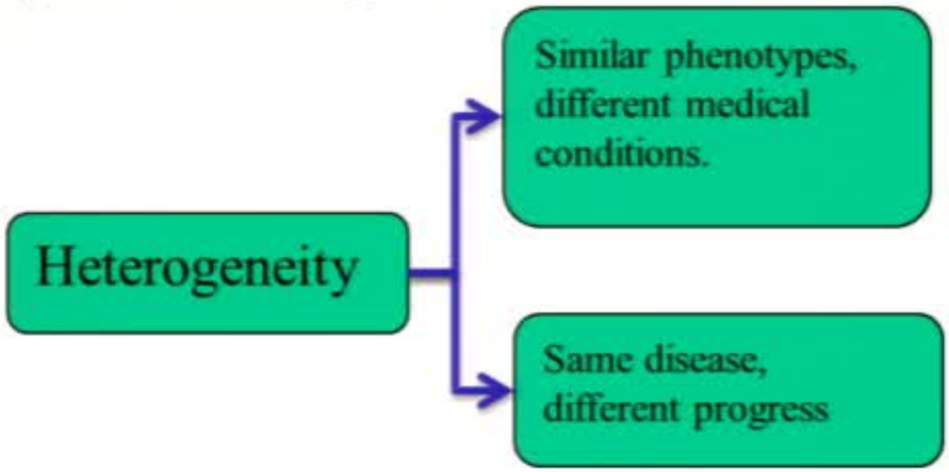


MICHIGAN

Personalized Medical Predictive Modeling

MICHIGAN

- Challenges: Heterogeneous nature of the patients



- How can we address this heterogeneity nature?

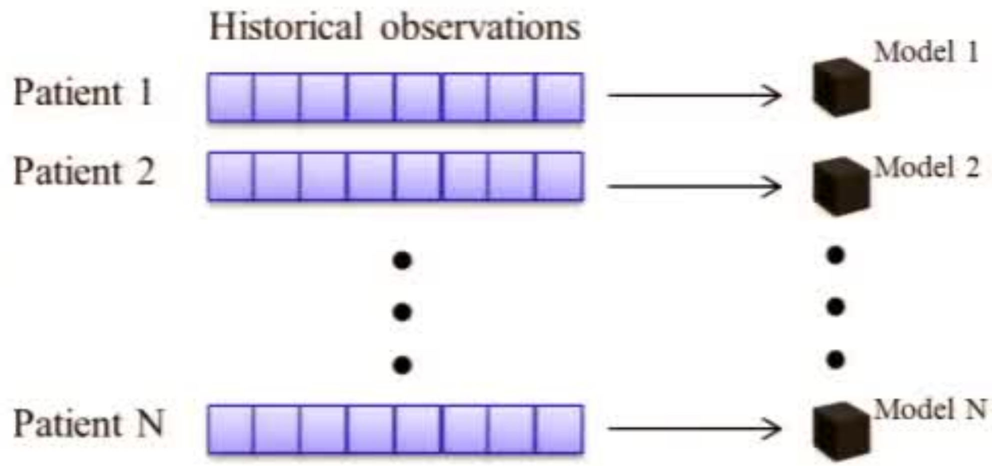
Personalized models

MICHIGAN

Personalized Medical Predictive Modeling

MICHIGAN

- A simple personalized modeling: learning independently



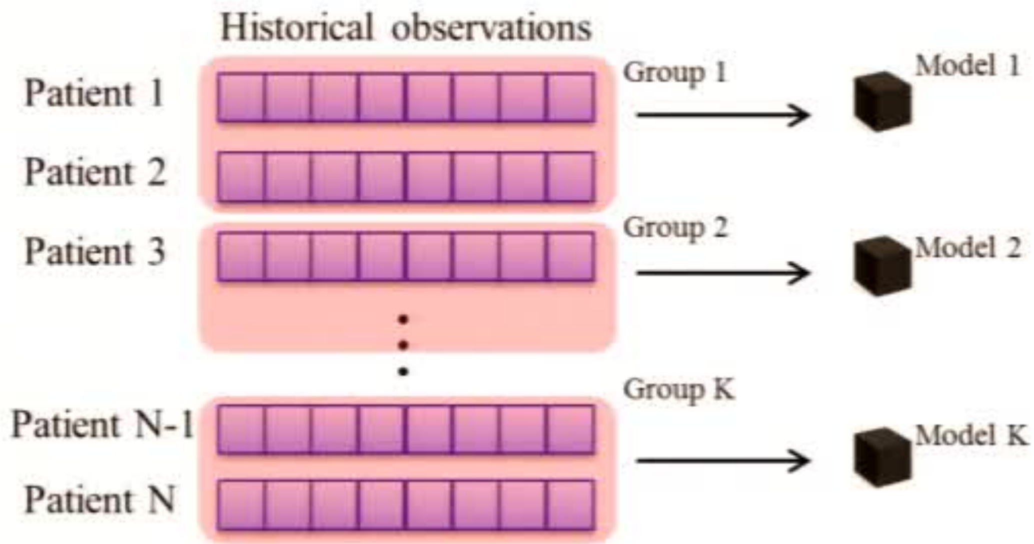
- Pros:
  - The idea is simple and straightforward.
- Cons:
  - It is not efficient in terms of time and space complexity when task number is large
  - The models will be likely to severely overfit the data due to the limited amount of training data and result in poor generalization performance

MICHIGAN

Personalized Medical Predictive Modeling

MICHIGAN

- A two-stage modeling method: grouping + modeling on each group



Examples:
Locally weighted learning
Localized support vector machine
Clustering + MTL

- Pros:
  - The number of models built are reduced
  - The number of training data is increased in each model building process
- Cons:
  - Grouping and model building are separate procedures, which could reach into suboptimal results
  - Building models for each group independently does not utilize potential valuable information from patients in other groups.
  - The models are not fully personalized as the patients within a group will have the same model

MICHIGAN

Personalized Medical Predictive Modeling

MICHIGAN

- Motivation:
  - **Personalized modeling**: build one model for each of the patient
  - **Patients grouping**: find hidden heterogeneity base models between patients
- Contribution:
  - we proposed a novel approach called FactORized MUlti-task LeArning model (FORMULA)
  - FORMULA extracts the base models of the patients and uses a linear combination of these models as the personalized model of a patient
  - FORMULA employs a sparse matrix factorization formulation to perform base model selection for each patient and feature selection for each base model.
  - We evaluated FORMULA on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The experimental results show the superiority of FORMULA over other baseline methods

Outline

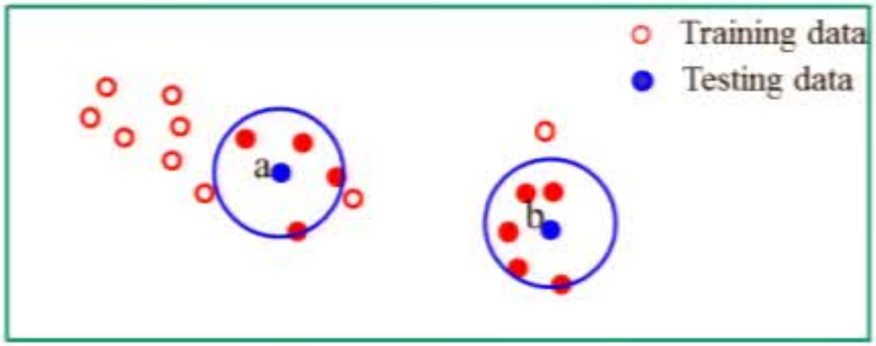
MICHIGAN

- Personalized medical predictive modeling
- Related works
- The proposed FORMULA Framework
- Experiments and Results
- Conclusion

Related works

MICHIGAN

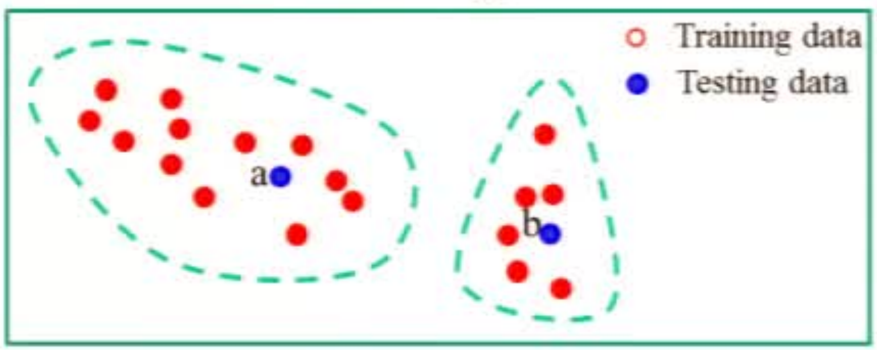
Locally weighted learning: lazy learning, building a model for each of the test data point



Drawbacks:

- Build a model for each test data points

Localized SVM: building a model for each group of the data point



Drawbacks:

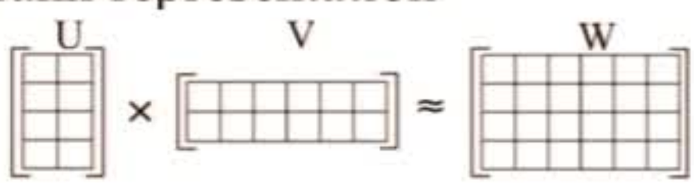
- Need all testing points in advance
- Not a fully personalized modeling procedure

Related works

MICHIGAN

Multi-task Learning: learning models for multiple tasks simultaneously by considering the information shared between tasks

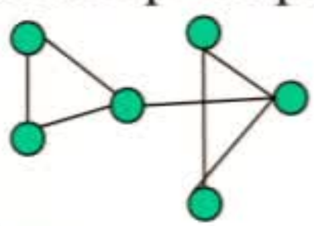
- Information sharing schemes:
  - Low-rank representation



Constraint

$$W = UV$$

  - Relationship Graph



Regularization

$$\sum_{i,j} A_{i,j} \|\mathbf{w}_i - \mathbf{w}_j\|_2^2$$

- Drawbacks:
  - Tasks are defined in advance.
  - It can be only used in two-stage modeling method as Clustering + MTL.

Outline

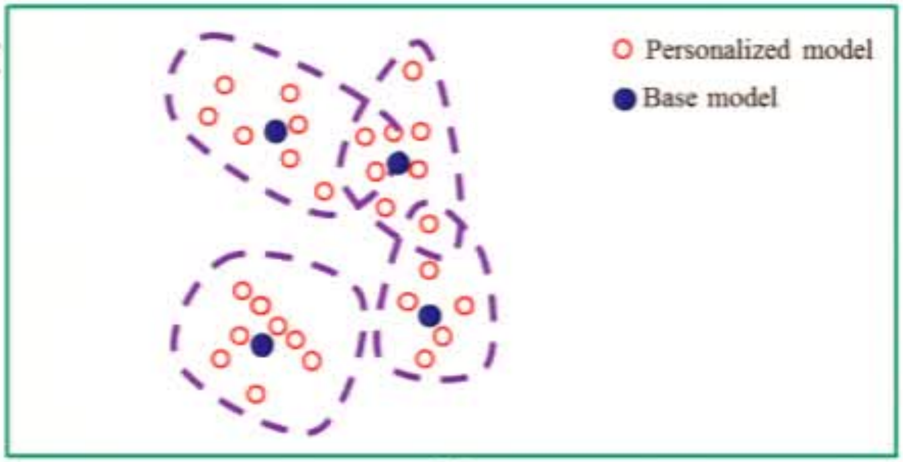
MICHIGAN

- Personalized medical predictive modeling
- Related works
- The proposed FORMULA Framework
- Experiments and Results
- Conclusion

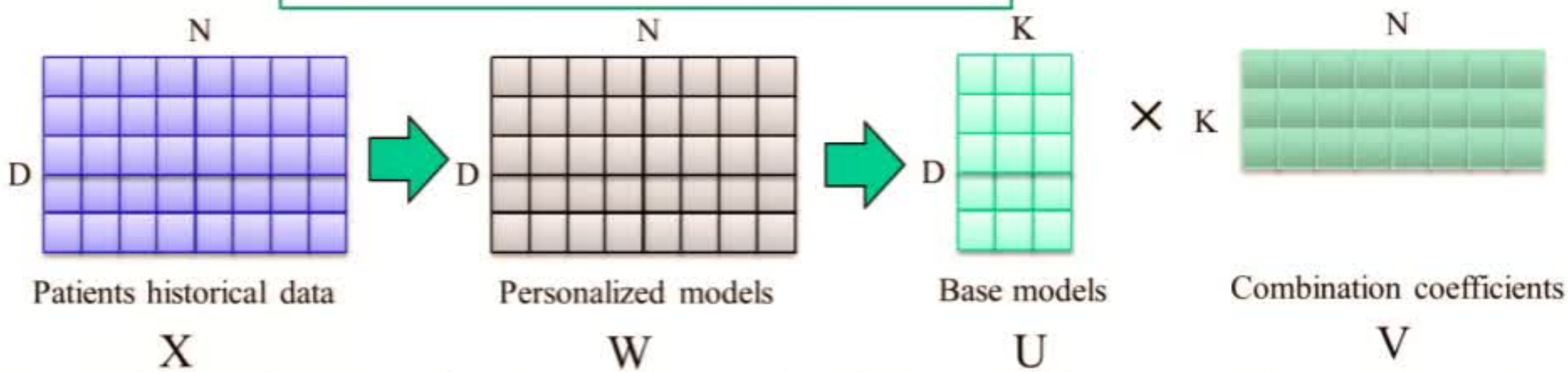
FORMULA

MICHIGAN

Basic idea:



Note: Grouping on models instead of data points



In order to learn robust and meaningful models, we enforce sparsity on U and V

FORMULA

MICHIGAN

Local Smoothness and Recovery:

- We assume the models for patients with similar phenotypes should be close to each other
- Such a model smoothness criterion is helpful to infer the personalized model for a test patient by assuming it is similar to the weighted average of the personalized models for its neighbors

FORMULA

Objective function:

Squared loss function

Sparsity on combination coefficients

$$\min_{W,U,V} \frac{1}{2} \boxed{\sum_i^N (y_i - \mathbf{w}_i^T \mathbf{x}_i)^2} + \lambda_1 \boxed{\|V\|_1}$$

$$+ \lambda_2 \boxed{\|U\|_1} + \frac{\lambda_3}{2} \boxed{\|W - WL\|_F^2}$$

$$\text{s.t. } V \succeq 0, \boxed{W = UV}$$

Sparsity on base models

Local Smoothness and Recovery

Task grouping and modeling

Outline

MICHIGAN

- Personalized medical predictive modeling
- Related works
- The proposed FORMULA Framework
- Experiments and Results
- Conclusion

MICHIGAN

Experiments and Results

MICHIGAN

Experiment setting:

- Data source: Alzheimer's disease Neuroimaging Initiative (ADNI) database
- Dataset: 10 datasets are created based on the elapsed time since the initial hospital visit

Table 1: Dataset size of ADAS-Cog and MMSE

	M06	M12	M24	M36	M48
ADAS-Cog	648	638	564	377	85
MMSE	648	642	569	389	87

- Baselines
  - SM: Single model with Ridge regression
  - CSTR: Clustering + single task model with Ridge regression
  - CSTL: Clustering + single task model with Lasso regression
  - CSL: Clustering + sparse low-rank MTL
  - CMR: Clustering + mean regularized MTL

MICHIGAN

Experiments and Results

MICHIGAN

Results:

Table 2: Comparison the MSE between FORMULA and baseline methods

MSE	ADAS-Cog					MMSE				
	M06	M12	M24	M36	M48	M06	M12	M24	M36	M48
SM	0.497	0.553	**0.758**	**0.893**	2.153	0.237	0.331	0.420	0.536	1.202
CSTR	0.469	0.637	0.879	0.985	1.543	0.246	0.329	0.392	0.493	0.747
CSTL	0.448	0.558	0.817	0.907	**1.416**	0.205	**0.264**	0.287	0.364	0.954
CSL	0.555	0.695	0.869	0.981	1.524	0.247	0.335	0.383	0.470	0.821
CMR	0.545	0.625	0.860	0.978	1.524	0.242	0.319	0.362	0.420	0.786
FORMULA	**0.424**	**0.545**	0.764	0.899	1.520	**0.197**	0.267	**0.262**	**0.347**	**0.673**

Remarks:

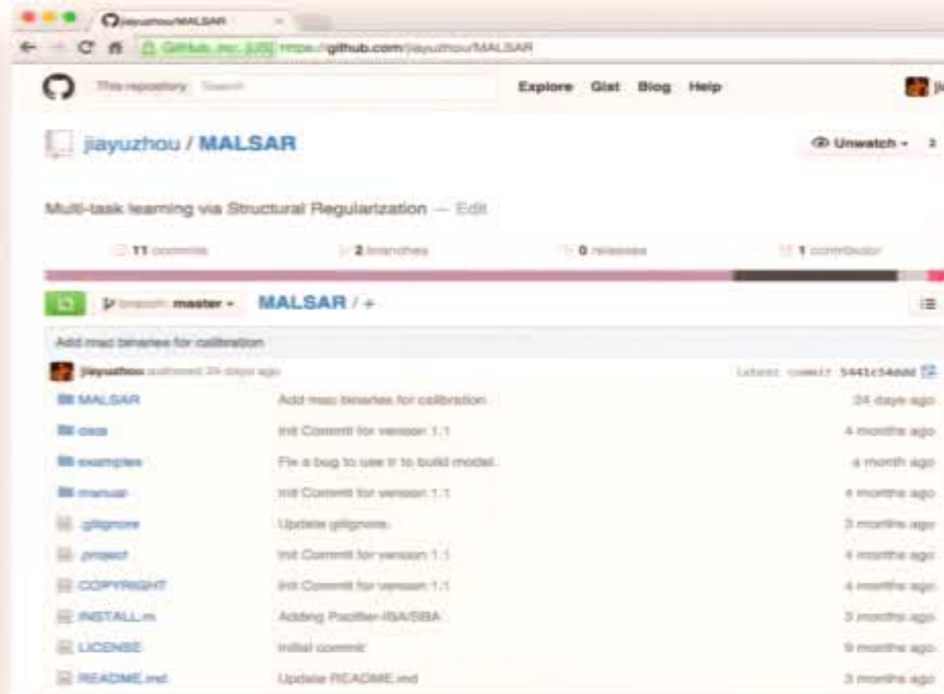
- FORMULA outperforms all other baselines in 6 out of 10 datasets
- FORMULA is consistently one of the top two algorithms for all the datasets
- FORMULA vs CSL/CMR: learning the groups and models simultaneously helps to improve the performance.

MICHIGAN

MALSAR

MULTI-TASK LEARNING VIA STRUCTURAL REGULARIZATION

MICHIGAN



- ⇨ **Efficient solvers of many state-of-the-art multi-task learning formulations;**
- ⇨ **Over 40 research works using MALSAR are published in KDD, NIPS, TPAMI, ICCV, ICDM, ICIP, COLING, MICCAI, ACM-MM, etc;**
- ⇨ **Used as course material to analyze compound profiling in the Strasbourg Summer School in France**

Outline

MICHIGAN

- Personalized medical predictive modeling
- Related works
- The proposed FORMULA Framework
- Experimental Results
- Conclusion

Conclusion

MICHIGAN

- we propose a novel FactORized MUlti-task LeArning model (FORMULA) to learn low-rank personalized models, leveraging the shared information among patients.
- FORMULA employs a sparse matrix factorization formulation to perform base model selection for each patient and feature selection for each base model.
- Our experimental results on the Alzheimer's Disease Neuroimaging Initiative (ADNI) data set suggest that the proposed approach is superior than several baseline method.