A recommender system for antimicrobial resistance

1st Sarath Swaminathan
Applied and Industrial Genomics
IBM Almaden Research Center
San Jose CA, USA
Sarath.Swaminathan@ibm.com

2nd Dmitry Yu. Zubarev
Accelerated Materials Discovery
IBM Almaden Research Center
San Jose CA, USA
dmitry.zubarev@ibm.com

3rd Mark Kunitomi
Applied and Industrial Genomics
IBM Almaden Research Center
San Jose CA, USA
Mark.Kunitomi@ibm.com

***Abstract*—We have developed a method for prediction of antimicrobial resistance by representing the antibiogram and associated metadata in the matrix form, determining parameter of matrix factorization, i.e., the number of the latent factors, carrying out matrix factorization of the initial association matrix and construction of the new association matrix containing predictions of the associations between multiple samples antibiogram and associated metadata.**

***Index Terms*—antimicrobial resistance, antibiogram, metadata, recommender, non-negative matrix factrorization**

I. Introduction

The surge of antibiotic resistance has made multi drug-resistant (MDR) bacterial infection a serious global threat. In 2014, it was estimated that by 2050, 10 million people will die every year from antimicrobial resistant (AMR) bacterial infection. Properly determining which antibiotics a particular microbe is resistant and susceptible to can be the difference between life and death. Therefore, understanding the features that are predictive of the antibiotics that a particular bacterium is susceptible and resistant to has, and will continue to become, a problem with significant impact on worldwide infection prognosis and treatment. Highlighting the significance of the threat of antimicrobial resistance, is the acceleration of the accumulation of data related to bacterial genomics and antibiotic resistance. As this rate increases, we will need to develop methods to rapidly aggregate, clean, and analyze the data. To this end, machine learning approaches to predicting AMR have the potential to scale with data and provide insights into this problem. Hospitals and governmental agencies periodically run tests to determine which antibiotics populations of bacteria are resistant/susceptible to. The overall profile of antimicrobial susceptibility testing results for each isolate is referred to as an antibiogram. We use the following definitions for the purposes of this contribution: *antibiogram* is a profile of antimicrobial susceptibility testing results of an isolate of a microorganism to various antimicrobial drugs; *antibiogram metadata* is the information associated with an antibiogram that describes its collection, including sequencing results, taxonomic classification, geographical region, source, date, etc.

Antibiograms can be used to guide a clinician and/or pharmacist in selecting the best empiric antimicrobial treatment in the event of pending microbiology culture and susceptibility results. Rapid assessment of the proper treatment for a bacterial infection can have drastic effects on the patient outcome. For example, patients with typhoid fever without timely and appropriate treatment was estimated to have a 30 percent mortality rate, whereas those that receive specific therapy, the rate reduced to 0.5 percent [1]. However, the method of producing an antibiogram requires culturing of the infecting organism and takes days to get results.

Previously produced antibiogram have the potential for use as tools for detecting and monitoring trends in antimicrobial resistance as well as prediction of unknown features of current infections. When antimicrobial susceptibility testing data are summarized cumulatively for a hospital, healthcare system over time, trends in resistance can be identified and investigated.

However, datasets that are aggregated from many sources, especially for those with a large number of potential measurements, the data can be sparse. This is because each source may have only collected data for a subset of the features. This is particularly true of antibiotic resistance data, known as antibiograms, that are aggregated by the National Center for Biotechnology (NCBI).

Each source within the set of antibiograms has performed tests to examine the susceptibility/resistance of an isolate of bacteria to a number of different antibiotics. Since there are many antibiotics to test, no individual entry has tested their particular bacterial isolate against every antibiotic. The lack of observations for many antibiotics for many bacterial isolates limits the descriptive, and therefore, predictive power of machine learning approaches to classify these isolates. In addition, the predictions themselves may be an endpoint prediction whereby making a limited set of observations provides statistically supported predictions of features not empirically observed.

In order to better understand trends in resistance as well as predict antibiotic resistance in clinical or industrial settings for newly observed microbial isolates, we need to be able to unify from various sources.
Matrix Factorization (MF) [2] is a form of latent factor

978-1-7281-1867-3/19/$31.00 ©2019 IEEE

analysis that enables link prediction in the context of network completion problem. This motivates broad application of matrix factorization in recommender systems that generate hypotheses about associations between objects, such as viewers and movies, shoppers and products, etc., via linear decomposition of the full-rank observation matrix into low-rank components. The task on hand - hypothesis generation regarding association between a sample and an antibiotic - falls in the same category and should be amenable to link prediction via MF or any appropriate approach. [3]
As a form of latent factor analysis, MF has been used for clustering tasks in microbiology. Reported applications of MF in this domain include detection of microbial communities in complex environments at the level of organism [4] and ocean habitats [5], dimensionality reduction in microbiome data [6], and phenotype assignment [7]. The method has not been applied to the prediction of antibiograms and their associated metadata.

Although matrix factorization has been used to successfully classify biological data, to date, no one has shown its application for antibiotic resistance, antibiograms, or other drug-microbe interactions.

RESULTS

Data Characterization

NARMS has generated human isolate data from 1996 to 2015 including metadata such as the classification, antibiotic susceptibility results, site of isolation, year of specimen collection, region, and age category (see Fig. 1 for the info-graphics of the dataset used in this study). Four genera are monitored by CDC's branch of NARMS: *Salmonella*, *Campylobacter*, *Shigella*, and *Escherichia*, which in total covers 8 species and over 400 serovars of *Salmonella*.

Each of these features or combination of them has the potential to reveal patterns of susceptibility/resistance for the antibiotics tested.

Region - Isolate collection is biased towards the Eastern and South Eastern regions of the United States.
Genus - Monitoring is heavily biased towards *Salmonella* and *Campylobacter*.
Age range - For each genus there is a moderate bias towards isolates towards younger patients for *Salmonella*, *Shigella*, and *Escherichia*, but not for *Campylobacter*. It is also clear that the majority of *Shigella* cases skew towards having resistance to at least one tested antibiotic, while *Escherichia* cases skew towards complete susceptibility.
Isolation source - For many of the isolates, the source of isolation is included. Here we show only four of the listed sources (Stool, Urine, Blood, and Wounds) as there were very few cases of other classified metadata in the dataset.
Resistance Phenotype - Reistance phenotye of an antibiotic is represented using three states - R (Resistant), S (Susceptible) and X (Inconclusive). For matrix factorization, we are only considering R and S entries in the data.

Data Visualization

For the purposes of MF analysis the data are represented as an adjacency matrix between isolates and antibiotic; Figure 2 shows data visualization as bi-partite network defined by this adjacency matrix. In this network antibiotics constitute one partition (dark blue nodes), the isolates labeled according to the assigned genus constitute another partition (green, light blue, orange, and red nodes). For the pair isolate-antibiotic the corresponding nodes are connected with a link if the isolate is susceptible to the antibiotic (please see Data Charatcerization section above for the mapping between NARMS CDC phenotype labels and "susceptible"/"resistant" phenotype labels used in computational experiments).

The force-based layout used to generate Figure 2 gives some impression about the clustering of the data; it is further elucidated by community analysis (Fig. 2B). Among three identified communities, Cluster I is composed primarily of Campylobacter (Campylobacter 71%, Salmonella 14%, Shigella 14%), Cluster II is primarily Salmonella (Salmonella 79%, Shigella 20%), and Cluster III is primarily Salmonella (Salmonella 82%, Shigella 9%, Escherichia 9%). The association of antibiotics with clusters is as follows: Cluster I - AZM,CIP,CLI,ERY,FFN,GEN,NAL,TEL; Cluster II - ATM,CAZ,CTX,FEP,IMI,PTZ; and Cluster III - AMI,AMP,AUG,AXO,CEP,CHL,COT,FIS,FOX,CAM, SMX,STR,TET,TIO.

Matrix Factorization Experiments

We benchmarked performance of several MF algorithms (see Methods section) in prediction of antimicrobial resistance of an isolate. A large number of MF algorithms are available and for benchmarking we selected four widely used algorithms. These MF algorithms are selected such that each of them have distincitive features like unconstarined (ALS), nonnegativity (NMF), implicit feedback (ALS-Implicit) and ranking based (MF-WARP).

The hypothesis generation task solved by MF-based recommender was formulated as follows. For a given uncharacterized pair isolate-antibiotic we aimed to predict its class as "resistant" or "susceptible". To this end we transformed the initial NAMRS dataset into two adjacency matrices between isolates and antibiotics. One matrix ("susceptile" class) has high weightage for susceptible entries and low weightage for resistant enries; the other matrix ("resistant" class) has high weightage for resistant entries and low weightage for susceptible entries.

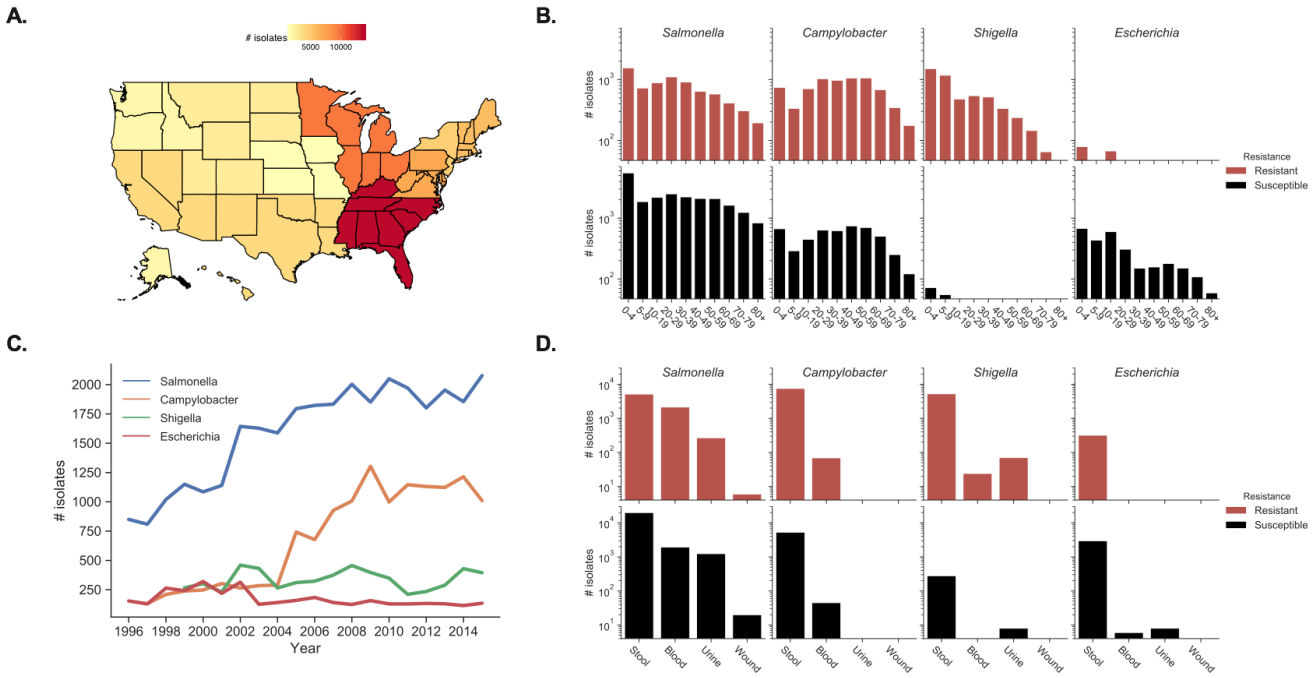

Fig. 1. Characterization of NARMS CDC patient data.(A) The number of isolates per region in the USA. (B) The number of isolates tested per genus over time. (C) Distribution of age-ranges of patients for the isolates. (D) Distribution of isolation sources.

The adjacency matrix for susceptible class was constructed in the following manner:

- ALS/ALS Implicit/MF-WARP: 1 (susceptile), -1 (resistant) and 0 (unknown);
- NMF: 5 (susceptible), 1 (resistant) and 0 (unknown).

The adjacency matrix for resistant class was constructed in the following manner:

- ALS/ALS Implicit/MF-WARP: 1 (resistant), -1 (susceptible) and 0 (unknown);
- NMF: 5 (resistant), 1 (susceptible) and 0 (unknown).

These association matrices were subject to MF analysis separately, so that for a given pair isolate-antibiotic one MF model generated hypothesis that its label is "suscpetible" and the other model generated the hypothesis that the label is "resistant". This approach also proved helpful in gaining insights into the severity of the class imabalance problem in NAMRS data: 92% of the available data are in "susceptible" class and only 8% are in "resistant" class. Since the data is imbalanced with very few resistant antibiotics, this approach helps to address underfitting and predict antimicrobial resistance with better accuracy. An MF model learned for predicting susceptible antibiotics need not be the best one for predicting resistant antibiotics. Also ranking based algorithms like MF-WARP works by improving the prediction accuracy for one class. Separate association matrix and models for predicting resistant and susceptible antibiotics makes more sense for such algorithms.

5-fold cross validation is used for parameter selection. 5-fold cross validation is performed by randomly partitioning known isolate-antibiotic pairs into five folds of equal size. After partitioning, MF model is trained on four folds and tested on the remaining one fold. This is repeated five times such that each of the folds will be used exactly one time for testing. 5-fold cross validation is then repeated three times on different random partitions.

Cold Start: MF requires some pre-existing information about isolate resistant/susceptible phenotype in order to predict new phenotypes. The situation when there is no such information available - no antibiotics have been tested against particular isolate - is referred to as a *cold start problem*. In order to resolve this problem we inferred phenotypes of uncharacterized isolates from the phenotypes of the most similar isolates. Isolate similarity was evaluated from the metadata, such as genus, species, serotype, age group, geographic region, and year: available metadata were vectorized using one-hot encoding, distances between isolates were evaluated as Euclidean distances, and the *predictions* for the phenotype of the isolates causing *cold start problem* was generated as a linear combination of the prediction of ten nearest isolates. *Cold start experiments* were performed based on the best performing matrix factorization model from previous experiment with 5-fold cross validation.

All-genera vs Genus-specific antibiograms: It is unclear how the presence of multiple genera in the compiled antibiogram

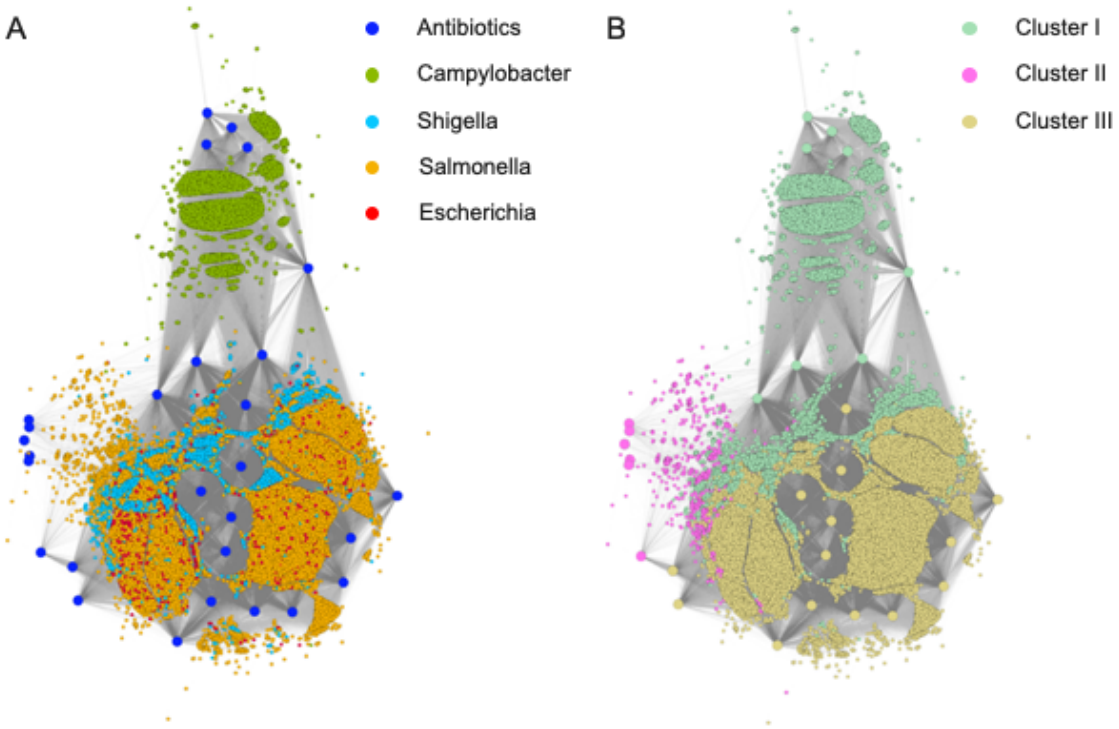

Fig. 2. AMR data used in this study represented and analyzed a bipartite network. (A) Antibiotics form the first partition, they are represented as large blue nodes; isolates form the second partition, they are represented as small nodes colored according to the sample Genus. If the sample shows susceptibility to the antibiotic, their respective nodes are connected with a link. (B) Modular structure of the network; nodes are colored according to their community assignment.

impacts predictions of antibiotic resistance produced by the recommender system. In order to address this question we prepared genus-specific datasets and benchmarked recommender system performance in all-genera and genus-specific cases.

Performance of various matrix factorization approaches for predicting susceptible and resistant antibiotics is summarized in Figure 3. ALS Implicit and NMF performed better than the other algorithms when predicting susceptible antibiotics with an AUCROC of 0.98, AUCPR of 0.96 and Precision @ 3 of 0.94. MF-WARP gave slightly less accuracy when prediciting susceptible antibiotics with AUCROC (0.96), AUCPR (0.92) and Precision @ 3 (0.88). When predicting resistant antibiotics, MF-WARP clearly outperform other algorithms with AUCROC of 0.95, AUCPR of 0.86 and Precision @ 3 of 0.81. And overall MF-WARP gives better result for predicting both susceptible and resistant antibiotics than other algorithms.

Cold start results based on MF-WARP is also shown in Figure 3. Cold start results are comparable with MF-WARP results for predicting susceptible and resistant antibiotics.

We further evaluated MF-WARP performance in all-genera and genus-specific recommendation tasks. Performance metrics are shown in Figure 4. Genus-specific performance was higher than all-genera performance for the prediction of both susceptible and resistant antibiotics in case of all four genera.

Discussion

Predicting ”resistant” phenotype is more important than ”susceptible” phenotype from both statistical and clinical perspectives. The former is the case because the data is heavily skewed towards susceptible phenotype. The latter is the case because making a false negative prediction about resistivity of an isolate to antibiotic is far more harmful than making a false positive prediction about susceptibility of the isolate, given larger fraction of susceptible isolates in the dataset.

Benchmarking results show high area under PR curve (AUPR) in prediction of the phenotype for both ”susceptible” and ”resistant” phenotypes. Among 4 versions of MF, MF-WARP showed most consistent results for both classes. Considering significant class imbalance in the data between phenotypes, it is highly encouraging that algorithm as simple as MF showed little sensitivity to it.

Another type of class imbalance is associated with taxonomic identity of isolates: the largest represented genus, Salmonella, comprises 58.7% of the isolates, whereas the smallest represented genus, Escherichia, comprises 6.3%.
In order to explore how the composition of the data, in regards

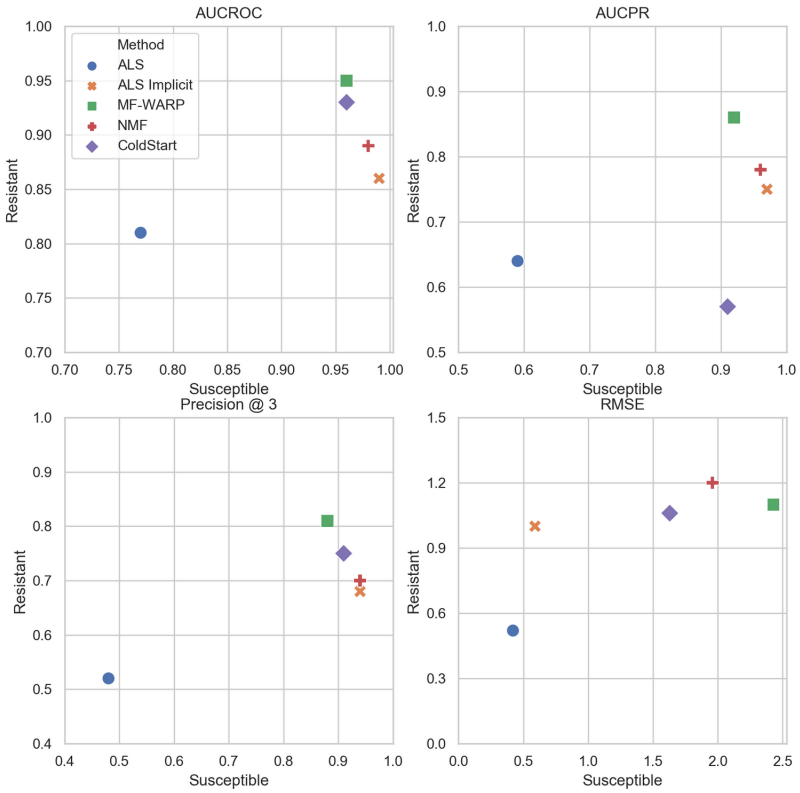

Fig. 3. Area under ROC (AUCROC) and PR (AUCPR) curves, Precision@3 and RMSE results for various matrix factorization approaches on all-genera antibiograms; prediction of the resistant and susceptible classes is performed separately.

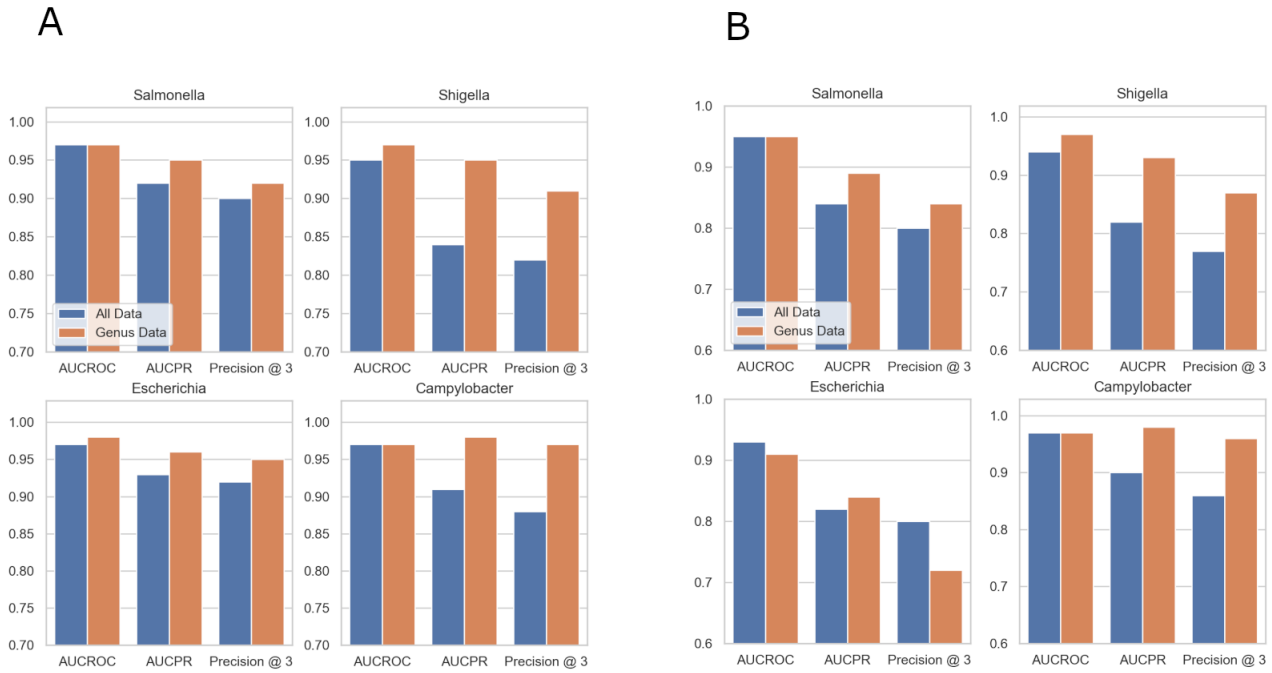

Fig. 4. Area under ROC and PR curve and Precision @ 3 results for predicting susceptible (panel A) and resistant (panel B) antibiotics with MF-WARP on the genus only and complete data

to genera included, affects the performance of prediction we performed MF on the combined dataset in comparison to each Genus separately. We found that genus specific data outperforms combined data, particularly for Shigella and Campylobacter, in the prediction of "resistant" phenotype. This is result shows that the size of the dataset, in this case restricted to individual genera, is not a driver for the MF performance because Shigella and Campylobacter are not the most abundant classes. A possible explanation for Campylobacter is its distinct antibiotic profile an observation further supported by the distinct clustering of Campylobacter in the bipartite network.

METHODS

Data

Data was downloaded from the CDC NARMS Now website: https://wwwn.cdc.gov/narmsnow/

Plotting

Tabulations of the NARMS data were performed in python using the pandas package. All plots were made with using either R (libraries: ggplot2 (https://cran.r-project.org/package=ggplot2) and fiftystater (https://CRAN.R-project.org/package=fiftystater)) or python (packages: matplotlib (https://matplotlib.org/) and Seaborn (https://seaborn.pydata.org/)). Construction of the network representation of the NARMS data was performed using python igraph package. Network visualization and modularity analysis were performed using Gephi software (https://gephi.org/)).

Matrix Factorization (MF)

Matrix factorization is a class of algorithms widely used in recommendation systems for its scalability and accuracy. MF algorithms work by factorizing the data matrix into two latent feature matrices of low dimensions that can reconstruct the initial data matrix. For a data matrix representing the resistance/susceptibility of isolates against antibiotics, MF learns two latent feature matrices, one for isolates and the other for antibiotics. Each row of the isolates latent feature matrix is the feature vector, u_i, learned for the corresponding isolate i. And each row of the antibiotics latent feature matrix represents the feature vector, v_j, for antibiotic j. From these learned feature vectors, resistance phenotype of an isolate i against an antibiotic j can be predicted by taking dot product of the respective feature vectors, $\hat{r}_{ij} = u_i.v_j$.

In general, matrix factorization algorithms learn latent features based on the known resistance/susceptible phenotypes present in the data matrix and doesn't require any additional metadata of the isolates or antibiotics. The concept behind MF is that, there are certain properties of isolates and antibiotics that explain the interactions present in the data matrix and the latent features learned by MF represent them. These latent features for an isolate could be simple and interpretable features like genus or species of the bacteria to something much more complex like presence of certain genes. Similarly for an antibiotic the latent features could be the class of antibiotics to the presence of certain molecular attributes in that antibiotic.

For applying matrix factorization algorithms, let's represent the antimicrobial resistance data of m isolates and n antibiotics as a matrix, $R \in \mathbb{R}^{m \times n}$, where

$$R_{ij} = \begin{cases} 1, & \text{if isolate } i \text{ is susceptible against antibiotic } j \\ -1, & \text{if isolate } i \text{ is resistant against antibiotic } j \\ 0, & \text{unknown} \end{cases} \tag{1}$$

Matrix factorization learns latent feature vectors for isolates (u) and antibiotics (v) by minimizing following objective function:

$$J = \sum_{i,j \in R_K} (R_{ij} - u_i^T v_j)^2 + \lambda(\|u_i\|^2 + \|v_j\|^2) \tag{2}$$

Where R_K is the pair of isolates and antibiotics with resistance/susceptable phenotype known, and λ is the regularization parameter used to avoid overfitting while learning latent feature vectors. Different matrix factorization algorithms are defined based on the objective function, restrictions applied on feature vectors and learning algorithms used for optimization.

In this work, we're using following four widely used matrix factorization algorithms:

Matrix Factorization with Alternating Least Squares (ALS): Basic matrix factorization with objective function defined in equation 2 and alternating least squares approach used for optimization [8].

Matrix Factorization with Implicit Feedback (ALS Implicit): Matrix factorization algorithms for recommendation systems mostly work on explicit feedback data available like the ratings users give to movies. Instead of this, MF with implicit feedback [9] factorizes implicit feedback derived from the explicit data along with a confidence score obtained from the explicit feedback by minimizing following objective function:

$$J = \sum_{i,j} c_{ij}(p_{ij} - u_i^T v_j)^2 + \lambda(\sum_i \|u_i\|^2 + \sum_j \|v_j\|^2) \tag{3}$$

Where p_{ij} is the implicit feedback and c_{ij} is the confidence score.

Nonnegative Matrix Factorization (NMF): NMF [10] is a matrix factorization algorithm which has a constraint that the data matrix and latent feature vectors are nonnegative.

Matrix Factorization with Weighted Approximate-Rank Pairwise loss (MF-WARP): The loss functions in traditional MF algorithms doesnt focus on the relative order between the predicted value of two antibiotics against an isolate. Weighted Approximate-Rank Pairwise (WARP) [11]

loss function learns latent factors such that the relative order of predicted value between antibiotics against an isolate is preserved.

For ALS and ALS Implicit, we've used implementation from Apache Spark [12]. scikit-learn [13] is used for NMF and Lightfm [14] is used for MF-WARP.

Evaluation

To benchmark the performance of various matrix factorization algorithms, following widely used metrics are used:

Area Under ROC/PR curve: For each isolate in the test data, predict resistance/susceptible phenotype value for all antibiotics except the ones present in the training data. Label susceptible antibiotics present in test data as 1, and rest of the antibiotics as 0. From this, calculate Area Under ROC curve (AUROC) and Area Under PR curve (AUPR) for each isolates. Calculate mean and standard deviation of AUROC and AUPR values for all isolates in the test data.

AUCROC and AUCPR measures the accuracy of binary classification, and here these metrics will measure how well the model can classify susceptible antibiotic and resistant antibiotics. While both AUCROC and AUCPR provide similar information for class balanced data, AUCPR is more suitable for class imbalanced data. In our data, there only 8% of data is resistant. So AUCPR will be more useful than AUCROC while measuring the performance of resistant antibiotics prediction.

Precision @ k: For each isolate in the test data, antibiotics which are not present in the training data are ordered based on the predicted phenotype value. From this ordered antibiotics, top k antibiotics are selected and precision score is calculated by

$$Precision\ @\ k = \frac{No.\ of\ susceptible\ antibiotics\ in\ top\ k}{k} \tag{4}$$

Precision @ k measures the quality of top k recommendations. In this case, it measures how accurately a model can predict top k susceptible/resistant antibiotics.

Root Mean Squared Error (RMSE): For all pairs of isolates and antibiotics in the test data, predict phenotype value and calculate RMSE. RMSE measures how well the model can reconstruct the initial data matrix.

All these metrics are calculated for predicting susceptible and resistant antibiotics.

AUTHOR CONTRIBUTIONS STATEMENT

M.K. , S.S. and D.Z. conceived the experiment(s) and conducted the experiment(s). All authors reviewed the manuscript.

ADDITIONAL INFORMATION

Competing interests

The authors declare no competing interests.

REFERENCES

[1] M. Krishnan, A. Suresh, G. Sumathy, and R. Srivani, "Studies on antimicrobial susceptibility pattern of salmonella isolates from Chennai, India," *International Journal of Pharma and Bio Sciences*, vol. 2, pp. 435–442, Apr. 2011.

[2] R. Tandon and S. Sra, "Sparse nonnegative matrix approximation: new formulations and algorithms," *Max Planck Institute for Biological Cybernetics*, pp. 1–19, Sep. 2010.

[3] S. Zhang, L. Yao, and A. Sun, "Deep learning based recommender system: A survey and new perspectives," *CoRR*, vol. abs/1707.07435, 2017. [Online]. Available: http://arxiv.org/abs/1707.07435

[4] P. Yang, X. Su, L. Ou-Yang, H.-N. Chua, X.-L. Li, and K. Ning, "Microbial community pattern detection in human body habitats via ensemble clustering framework," *BMC Systems Biology*, vol. 8, no. Suppl 4, p. S7, 2014.

[5] X. Jiang, M. G. I. Langille, R. Y. Neches, M. Elliot, S. A. Levin, J. A. Eisen, J. S. Weitz, and J. Dushoff, "Functional biogeography of ocean microbes revealed through non-negative matrix factorization." *PloS one*, vol. 7, no. 9, p. e43866, 2012.

[6] Y. Cai, H. Gu, and T. Kenney, "Learning Microbial Community Structures with Supervised and Unsupervised Non-negative Matrix Factorization," *Microbiome*, vol. 5, no. 1, p. 349, Aug. 2017.

[7] M. Brbić, M. Piškorec, V. Vidulin, A. Kriško, T. Smuc, and F. Supek, "Phenotype Inference from Text and Genomic Data," *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Jan. 2017.

[8] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009. [Online]. Available: http://dx.doi.org/10.1109/MC.2009.263

[9] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*, ser. ICDM '08. Washington, DC, USA: IEEE Computer Society, 2008, pp. 263–272. [Online]. Available: http://dx.doi.org/10.1109/ICDM.2008.22

[10] A. Cichocki and A.-H. Phan, "Fast Local Algorithms for Large Scale Nonnegative Matrix and Tensor Factorizations," *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. 92, pp. 708–721, 2009.

[11] J. Weston, S. Bengio, and N. Usunier, "Wsabie: Scaling up to large vocabulary image annotation," in *Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI*, 2011.

[12] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M. J. Franklin, A. Ghodsi, J. Gonzalez, S. Shenker, and I. Stoica, "Apache spark: A unified engine for big data processing," *Commun. ACM*, vol. 59, no. 11, pp. 56–65, Oct. 2016. [Online]. Available: http://doi.acm.org/10.1145/2934664

[13] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[14] M. Kula, "Metadata embeddings for user and item cold-start recommendations," in *Proceedings of the 2nd Workshop on New Trends on Content-Based Recommender Systems co-located with 9th ACM Conference on Recommender Systems (RecSys 2015), Vienna, Austria, September 16-20, 2015.*, ser. CEUR Workshop Proceedings, T. Bogers and M. Koolen, Eds., vol. 1448. CEUR-WS.org, 2015, pp. 14–21. [Online]. Available: http://ceur-ws.org/Vol-1448/paper4.pdf