

A top-down view of a medical examination table. A patient's arm is extended, with a blue blood pressure cuff wrapped around the upper arm. A healthcare provider's hands are visible, one holding the patient's hand and the other near the cuff. A stethoscope lies on a clipboard to the left, which contains a 'REPORT SCHEDULE' form. A black coiled tube for the blood pressure monitor is visible on the right. The background is a light blue surface.

Exploring the Big Data Healthcare Transformation with Google Cloud

Team 8

Ryan Burger, Laura Cattaneo,
Allen Hu, and Jane Jian

Agenda

1. Key Takeaways

2. Overview of Healthcare in the Cloud

3. Demonstration

- De-identifying patient FHIR data
- Predicting probability of death due to medical reasons with BigQuery ML and AutoML



Key Takeaways



Data is transforming the healthcare industry and Google Cloud Platform has tools to help keep up



Unique healthcare data types, and patient privacy concerns make adoption a challenge



Google Cloud Platform presents a user-friendly and adoptable solution, benefitting both patients and healthcare providers



Overview of Healthcare in the Cloud



Healthcare is undergoing a transformation

Mayo Clinic selects Google as strategic partner for health care innovation, cloud computing

September 10, 2019

Cleveland Clinic, IBM partner; accelerating discovery in healthcare, life sciences

April 5, 2021

Microsoft and Providence St. Joseph Health announce strategic alliance to accelerate the future of care delivery

July 8, 2019 | Microsoft News Center

Research Highlight | Published: 21 January 2020

BREAST CANCER

AI outperforms radiologists in mammographic screening

David Killock 

Nature Reviews Clinical Oncology **17**, 134(2020) | [Cite this article](#)

OUTLOOK • 18 NOVEMBER 2020

Artificial intelligence is improving the detection of lung cancer

Machine learning systems for early detection could save lives.

Patients and providers benefit from cloud computing



Providers

- Lower operational costs and more personalized care
- Increased data access and data dimensions for analytics



Patients

- Access to personal data and increased interactive patient experience
- Increased participation and telemedicine capabilities

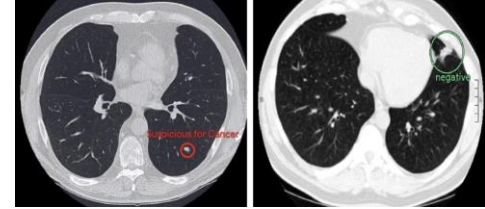


Healthcare industry uses unique data types



Fast Healthcare Interoperability Resources (FHIR)

- Data standard and API
- Share and understand data across applications or organizations



Digital Imaging and Communications in Medicine (DICOM)

- Standardized medical images
- Store and share medical images across hospitals, and image manufacturers

Technologies must meet specific requirements



HIPAA: Health Insurance Portability and Accountability Act

Patient de-identification with FHIR and DICOM data

Patient ID	First Name	Last Name	Weight
123	Ryan	Burger	3000 lbs
234	Laura	Cattaneo	1000 lbs
345	Jane	Jian	1000 lbs
567	Allen	Hu	2000 lbs



Demonstration





De-identifying FHIR data and using it in BigQuery

Code adapted from [Qwik Labs Demo](#)







Predicting probability of death due to medical reasons with BigQuery ML (BQML) and AutoML

Code adapted from [GCP Demo Github](#)





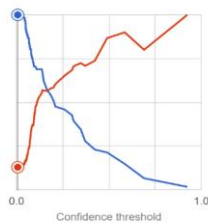
Code adapted from [GCP Demo Github](#)



Flexibility to train various models in BQML

Logistic Regression

Precision@1 and Recall@1 ?



ROC (Receiver Operating Characteristic) ?

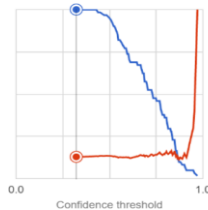


Area under curve: 0.844

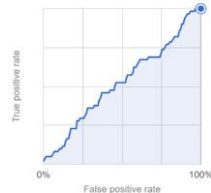
AUC: 0.844

XGBoosted Tree

Precision@1 and Recall@1 ?



ROC (Receiver Operating Characteristic) ?

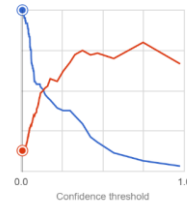


Area under curve: 0.523

AUC: 0.523

Deep Neural Network

Precision@1 and Recall@1 ?



ROC (Receiver Operating Characteristic) ?



Area under curve: 0.806

AUC: 0.806

Better performing Neural Network with AutoML



Multiple ML tools available on
Google Cloud Platform:

BQML

- Can train many kinds of models
- Hands on model tuning with moderate performance
- Low cost

AutoML

- Uses Neural architecture for specific problem
- Hands off ML with high performance
- Higher cost in GCP (\$20/hour)



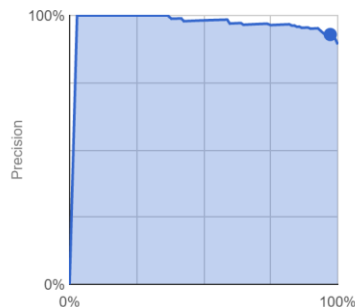
Better performing Neural Network with AutoML



Multiple ML tools available on
Google Cloud Platform:

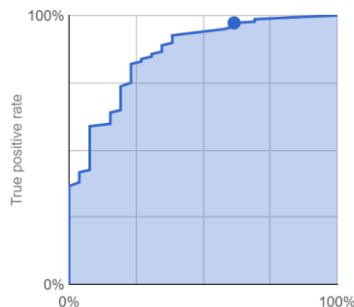


Neural network & overall performance improved



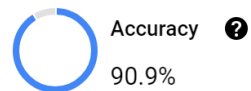
AUC: 0.979

PRC ?



AUC: 0.869

ROC ?



Binary classification model
Apr 12, 2021, 4:12:05 PM
Training cost: 1.811 node hours

Log loss ?
0.254

Feel free to contact us with any questions



Ryan Burger

burge218@umn.edu



Laura Cattaneo

catta008@umn.edu



Allen Hu

hu000345@umn.edu



Jane Jian

jian0450@umn.edu

For more technical details, check out our [GitHub](#)

Works Cited

- **Sources:**

- <https://github.com/synthetichealth/synthea/wiki>
- <https://towardsdatascience.com/choosing-between-tensorflow-keras-bigquery-ml-and-automl-natural-language-for-text-classification-6b1c9fc21013>
- <https://www.todaysmedicaldevelopments.com/video/cleveland-clinic-ibm-medical-healthcare-ai/>
- <https://news.microsoft.com/2019/07/08/microsoft-and-providence-st-joseph-health-announce-strategic-alliance-to-accelerate-the-future-of-care-delivery/>
- <https://newsnetwork.mayoclinic.org/discussion/mayo-clinic-selects-google-as-strategic-partner-for-health-care-innovation-cloud-computing/>
- <https://www.nature.com/articles/d41586-020-00847-2>
- <https://www.nature.com/articles/s41571-020-0329-7>
- <https://www.healthitoutcomes.com/doc/ways-cloud-computing-is-impacting-healthcare-0001>

- **Photos:**

- <https://innovacer.com/blogs/the-abc-of-fhir-reinventing-interoperability/>
- <https://venturebeat.com/2019/05/20/googles-lung-cancer-detection-ai-outperforms-6-human-radiologists/>