



Figure 5: The $\delta(c, \Delta)$ function for six different codes.

for healthcare using reverse time attention mechanism. In *Advances in Neural Information Processing Systems*. 3504–3512.

- [9] Edward Choi, Andy Schuetz, Walter F Stewart, and Jimeng Sun. 2016. Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association* 24, 2 (2016), 361–370.
- [10] Yujuan Feng, Xu Min, Ning Chen, Hu Chen, Xiaolei Xie, Haibo Wang, and Ting Chen. 2017. Patient outcome prediction via convolutional neural networks based on multi-granularity medical concept embedding. In *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 770–777.
- [11] Djordje Gligorijevic, Jelena Stojanovic, and Zoran Obradovic. 2016. Disease types discovery from a large database of inpatient records: A sepsis study. *Methods* 111 (2016), 45–55.
- [12] Yoni Halpern, Steven Hornig, Youngduck Choi, and David Sontag. 2016. Electronic medical record phenotyping using the anchor and learn framework. *Journal of the American Medical Informatics Association* 23, 4 (2016), 731–740.
- [13] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [14] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [15] Carrie N Klabunde, Arnold L Potosky, Julie M Legler, and Joan L Warren. 2000. Development of a comorbidity index using physician claims data. *Journal of clinical epidemiology* 53, 12 (2000), 1258–1267.
- [16] Harlan M Krumholz, Yun Wang, Jennifer A Mattern, Yongfei Wang, Lein Fang Han, Melvin J Ingber, Sheila Roman, and Sharon-Lise T Normand. 2006. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. *Circulation* 113, 13 (2006), 1693–1701.
- [17] Nathan Levitan, A Dowlati, SC Remick, HI Tahsildar, LD Sivinski, R Beyth, and AA Rimm. 1999. Rates of initial and recurrent thromboembolic disease among patients with malignancy versus those without malignancy. *Risk analysis using Medicare claims data. Medicine (Baltimore)* 78, 5 (1999), 285–91.
- [18] Zachary C Lipton, David C Kale, Charles Elkan, and Randall Wetzel. 2015. Learning to diagnose with LSTM recurrent neural networks. *arXiv preprint arXiv:1511.03677* (2015).
- [19] Fenglong Ma, Radha Chitta, Jing Zhou, Quanzeng You, Tong Sun, and Jing Gao. 2017. Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1903–1911.
- [20] Riccardo Miotto, Fei Wang, Shuang Wang, Xiaoqian Jiang, and Joel T Dudley. 2017. Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics* (2017).
- [21] Phuoc Nguyen, Truyen Tran, Nilmini Wickramasinghe, and Svetha Venkatesh. 2017. Deepcr: A Convolutional Net for Medical Records. *IEEE journal of biomedical and health informatics* 21, 1 (2017), 22–30.
- [22] Trang Pham, Truyen Tran, Dinh Phung, and Svetha Venkatesh. 2016. Deepcare: A deep dynamic memory model for predictive medicine. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 30–41.
- [23] Sanjay Purushotham, Chuizheng Meng, Zhengping Che, and Yan Liu. 2017. Benchmark of Deep Learning Models on Large Healthcare MIMIC Datasets. *arXiv preprint arXiv:1710.08531* (2017).
- [24] Alvin Rajkomar, Eyal Oren, Kai Chen, Andrew M Dai, Nissan Hajaj, Peter J Liu, Xiaobing Liu, Mimi Sun, Patrik Sundberg, Hector Yee, et al. 2018. Scalable and accurate deep learning for electronic health records. *arXiv preprint arXiv:1801.07860* (2018).
- [25] Narges Razavian, Jake Marcus, and David Sontag. 2016. Multi-task prediction of disease onsets from longitudinal laboratory tests. In *Machine Learning for Healthcare Conference*. 73–100.
- [26] Sebastian Schneeweiss, John D Seeger, Malcolm Maclure, Philip S Wang, Jerry Avorn, and Robert J Glynn. 2001. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *American journal of epidemiology* 154, 9 (2001), 854–864.
- [27] Mike Schuster and Kuldeep K Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* 45, 11 (1997), 2673–2681.
- [28] Ying Sha and May D Wang. 2017. Interpretable Predictions of Clinical Outcomes with An Attention-based Recurrent Neural Network. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 233–240.
- [29] Donald H Taylor Jr, Truls Østbye, Kenneth M Langa, David Weir, and Brenda L Plassman. 2009. The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited. *Journal of Alzheimer’s Disease* 17, 4 (2009), 807–815.
- [30] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*. 6000–6010.
- [31] Joan L Warren, Linda C Harlan, Angela Fahey, Beth A Virnig, Jean L Freeman, Carrie N Klabunde, Gregory S Cooper, and Kevin B Knopf. 2002. Utility of the SEER-Medicare data to identify chemotherapy use. *Medical care* 40, 8 (2002), IV–55.
- [32] Pranjul Yadav, Michael Steinbach, Vipin Kumar, and Gyorgy Simon. 2017. Mining Electronic Health Records: A Survey. *arXiv preprint arXiv:1702.03222* (2017).
- [33] Yan Yan, Elena Birman-Deych, Martha J Radford, David S Nilasena, and Brian F Gage. 2005. Comorbidity indices to predict mortality from Medicare data: results from the national registry of atrial fibrillation. *Medical care* 43, 11 (2005), 1073–1077.
- [34] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1480–1489.
- [35] Kaiping Zheng, Wei Wang, Jinyang Gao, Kee Yuan Ngiam, Beng Chin Ooi, and Wei Luen James Yip. 2017. Capturing Feature-Level Irregularity in Disease Progression Modeling. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. ACM, 1579–1588.