# Don't Predict Counterfactual Values, Predict Expected Values Instead

**Jeremiasz Wołosiuk**[1], **Maciej Świechowski**[2,3], **Jacek Mańdziuk**[3]

[1] Deepsolver
[2] QED Software
[3] Warsaw University of Technology
jeremi@deepsolver.com, maciej.swiechowski@qed.pl, mandziuk@mini.pw.edu.pl

## Bucketed losses

In the main body of the paper (Table 2), we present bucketed loss of DEVN and DCVN, in setting II. For both networks, one can observe a clear decline of both training and validation losses along with an increasing size of the networks. Furthermore, in each case, the values of both losses are close to each other. The above observations suggest that neural networks, in both DEVN and DCVN, learn their respective tasks effectively.

Qualitatively the same conclusions are valid for other learning settings which were not included in the main document due to space limits. These detailed results are presented in Tables 1- 5, for settings I-V, respectively.

## Comments on the zero-sum property (ZSP)

Texas Hold'em has a zero-sum property, i.e. for any terminal history $h_t$, $\sum\limits_{i \in P} u_i(h_t) = 0$, because a pot consists only of players' money and the gain of one player is a loss of all the others. This property translates into the fact, that a dot product of CFVs for both players and the respective ranges is equal to zero.

During DCVN training, ZSP can be enforced on the bucketed CFVs by means of an outer network which imposes that bucketed CFVs multiplied by bucketed ranges sum up to zero. After inverse bucketing, ZSP still holds because for each bucket, individual CFVs are a copy of bucketed CFV, and a bucketed range value is a sum of all individual range values.

When training DEVN, ZSP cannot be enforced directly on the bucketed EVs because the outer network would have to provide matchups for bucketed ranges, which is not feasible. However, after the inverse bucketing is performed and the matchups are calculated, ZSP can be easily enforced on the obtained CFVs, as an additional (last) step of the DEVN pipeline.

In this paper the impact of not enforcing ZSP on the playing strength of a Poker bot was not verified, as such an assessment requires having an access to a bot that implements the DEVN pipeline. We left this additional evaluation of DEVN as a further study.



## Visualization of CFVs bucketing process and encoding error calculation

The process of EVs bucketing and inverse bucketing (the DEVN pipeline) has been illustrated in Figure 1 in the main paper and discussed in detail throughout the text. DEVN modifies the DCVN pipeline implemented by top Poker bots (Moravčík et al. 2017; Zarick et al. 2020) by separating calculation of matchups from EVs estimation.

To facilitate direct comparison of both pipelines, the DCVN pipeline, which performs bucketing and inverse bucketing directly on CFVs, is presented in Figure 1.

## Experiments on neural networks without bucketing

To test if superior performance of DEVN over DCVN can indeed be attributed to a different usage of the card abstraction, we created two settings, in which the bucketing is not used. Alternatively, these settings can be regarded as using the "identity bucketing", where each hand goes into its own bucket, and additionally the network gets one-hot encoding vector of public cards as extended input. The vector has $52$ elements, one for each card in a deck, with a value of $1$ if the card is present on the board, and $0$, otherwise.

The experiments show that the use of DEVN introduces a relative improvement of $3.37 - 8.39\%$ over using DCVN, which indicates that prediction of EVs is a simpler problem than CFVs prediction.

Experiments were performed starting from 7-layer networks downto 3-layer ones. 2-layer networks were missing, because it was noted that already 3-layer network had significant errors and using smaller network wouldn't be feasible. The fact that neural networks without bucketing have significantly worse performance compared to the settings with the bucketing is attributed to the increased difficulty of the task, as using card abstraction provides substantial simplification of the problem.

The unbucketed loss gain relative to the 3-layer network, for each neural network size and for both DEVN and DCVN is illustrated in Figure 2. It can be seen that while in setting VI (referring to Dataset 1) DEVN gains more from adding layers, in setting VII (referring to Dataset 2) DEVN and DCVN have similar characteristics. Both settings are described in Table 6.

Table 1: Bucketed losses in setting I.

| Method | Training loss | Validation loss |
|---|---|---|
| EVs (2 layers) | 0.001080 | 0.001420 |
| EVs (3 layers) | 0.000887 | 0.001173 |
| EVs (5 layers) | 0.000747 | 0.001008 |
| EVs (7 layers) | 0.000707 | 0.000949 |
| CFVs (2 layers) | 0.000961 | 0.001211 |
| CFVs (3 layers) | 0.000801 | 0.001025 |
| CFVs (5 layers) | 0.000695 | 0.000898 |
| CFVs (7 layers) | 0.000659 | 0.000835 |

Table 2: Bucketed losses $\pm$ st.dev. in setting II.

| Method | Training loss | Validation loss |
|---|---|---|
| EVs (2 layers) | 0.000796 $\pm$ 3e-6 | 0.001270 $\pm$ 4e-6 |
| EVs (3 layers) | 0.000628 $\pm$ 8e-7 | 0.001007 $\pm$ 4e-6 |
| EVs (5 layers) | 0.000534 $\pm$ 2e-6 | 0.000830 $\pm$ 4e-6 |
| EVs (7 layers) | 0.000507 $\pm$ 3e-6 | 0.000780 $\pm$ 7e-7 |
| CFVs (2 layers) | 0.000746 $\pm$ 1e-6 | 0.001093 $\pm$ 5e-6 |
| CFVs (3 layers) | 0.000612 $\pm$ 1e-6 | 0.000891 $\pm$ 3e-6 |
| CFVs (5 layers) | 0.000540 $\pm$ 2e-7 | 0.000761 $\pm$ 9e-9 |
| CFVs (7 layers) | 0.000516 $\pm$ 2e-6 | 0.000726 $\pm$ 3e-6 |

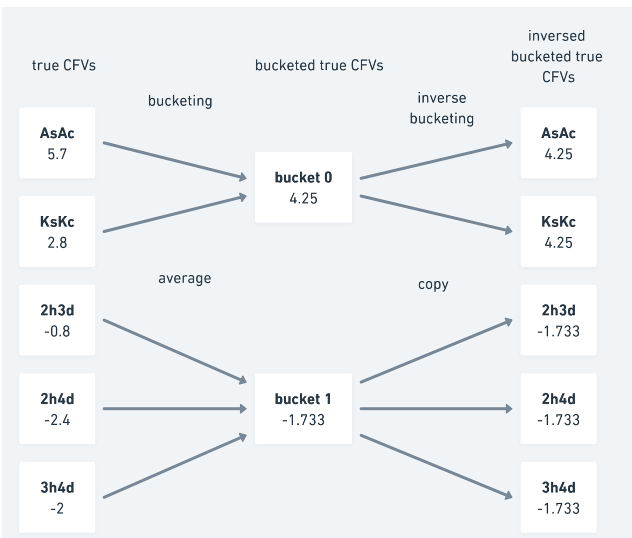


Figure 1: The processes of CFVs bucketing and inverse bucketing performed on a subset of hands. During the bucketing, CFVs for hands in a specific bucket are averaged. During inverse bucketing, the value for a hand assigned to a bucket is a copy of the bucket value. Presented values are random. $encoding\_error = \sqrt{(5.7-4.25)^2+\ldots+(-2-(-1.733))^2} = 2.36$.

Table 3: Bucketed losses in setting III.

| Method | Training loss | Validation loss |
|---|---|---|
| EVs (2 layers) | 0.000664 | 0.001245 |
| EVs (3 layers) | 0.000533 | 0.000997 |
| EVs (5 layers) | 0.000462 | 0.000820 |
| EVs (7 layers) | 0.000436 | 0.000760 |
| CFVs (2 layers) | 0.000626 | 0.001069 |
| CFVs (3 layers) | 0.000526 | 0.000872 |
| CFVs (5 layers) | 0.000471 | 0.000732 |
| CFVs (7 layers) | 0.000452 | 0.000695 |

Table 4: Bucketed losses in setting IV.

| Method | Training loss | Validation loss |
|---|---|---|
| EVs (2 layers) | 0.002800 | 0.004733 |
| EVs (3 layers) | 0.002272 | 0.004050 |
| EVs (5 layers) | 0.001927 | 0.003602 |
| EVs (7 layers) | 0.001864 | 0.003295 |
| CFVs (2 layers) | 0.002577 | 0.003975 |
| CFVs (3 layers) | 0.002191 | 0.003471 |
| CFVs (5 layers) | 0.001957 | 0.003141 |
| CFVs (7 layers) | 0.001865 | 0.003007 |

Table 5: Bucketed losses in setting V.

| Method | Training loss | Validation loss |
|---|---|---|
| EVs (2 layers) | 0.002353 | 0.004889 |
| EVs (3 layers) | 0.001935 | 0.004086 |
| EVs (5 layers) | 0.001668 | 0.003467 |
| EVs (7 layers) | 0.001589 | 0.003182 |
| CFVs (2 layers) | 0.002146 | 0.004075 |
| CFVs (3 layers) | 0.001847 | 0.003538 |
| CFVs (5 layers) | 0.001650 | 0.003077 |
| CFVs (7 layers) | 0.001589 | 0.002918 |

Table 6: Unbucketed loss of DCVN and DEVN for settings not using card abstraction. EE denotes the encoding error, which is equal to 0, because "identity bucketing" does not introduce any errors.

| Size | DCVN | DEVN | Rel. impr. |
|---|---|---|---|
| ***VI: Dataset 1; no bucketing; 1024 neurons per layer*** | | | |
| 3 | 5.309 | 5.046 | 4.95% |
| 5 | 4.150 | 3.801 | 8.39% |
| 7 | 3.808 | 3.497 | 8.16% |
| EE | 0 | 0 | – |
| ***VII: Dataset 2; no bucketing; 1024 neurons per layer*** | | | |
| 3 | 8.770 | 8.461 | 3.51% |
| 5 | 7.183 | 6.941 | 3.37% |
| 7 | 6.905 | 6.620 | 4.13% |
| EE | 0 | 0 | – |

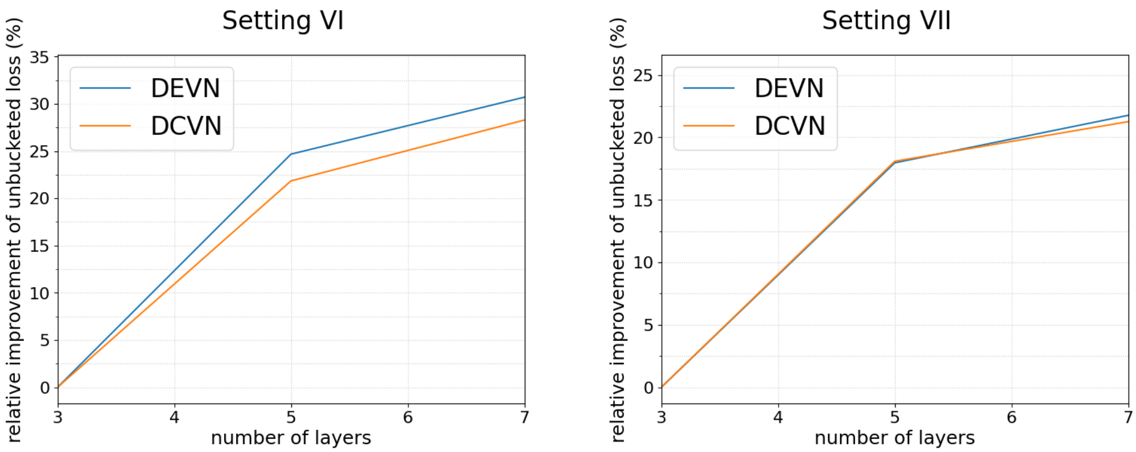


Figure 2: Unbucketed loss for DEVN and DCVN, for each neural network size, relative to the 3-layer network.

## References

Moravčík, M.; Schmid, M.; Burch, N.; Lisý, V.; Morrill, D.; Bard, N.; Davis, T.; Waugh, K.; Johanson, M.; and Bowling, M. 2017. DeepStack: Expert-level Artificial Intelligence in heads-up no-limit Poker. *Science*, 356(6337): 508–513.

Zarick, R.; Pellegrino, B.; Brown, N.; and Banister, C. 2020. Unlocking the Potential of Deep Counterfactual Value Networks. arXiv:2007.10442.