

## **Appendix for : A Closer Look At Feature Space Data Augmentation For Few-Shot Intent Classification**

**Varun Kumar, Hadrien Glaude, Cyprien de Lichy, William Campbell**

Amazon Alexa

Cambridge, MA, USA

{kuvrun, hglaude, cllichy, cmpw}@amazon.com

### **A FSI experiment results for all intents**

In all tables, individual columns represent FSI results for an intent, and *Overall Mean* column, provides average accuracy for all intents' FSI simulations. Please scroll to next page for results tables.

#	Method	Playlist	Restaurant	Weather	Music	Book	Work	Event	Overall Mean
No Augmentation		82.63(5.11)	87.86(3.53)	84.51(1.3)	88.07(2.37)	96.81(2.94)	85.14(1.53)	87.19(3.31)	87.46(2.87)
100	UPSAMPLE	92.24(2.96)	97.7(0.67)	96.44(0.75)	94.57(1.1)	97.96(0.82)	89.61(3.01)	91.26(2.35)	94.26(1.66)
	PERTURB	<b>93.09</b> (2.55)	97.41(0.92)	96.07(1.35)	94.39(1.13)	97.86(0.93)	89.36(2.76)	91.09(2.53)	94.18(1.74)
	CVAE	92.4(3.66)	97.47(0.67)	96.49(1.07)	94.36(1.26)	97.71(1.1)	89.1(2.79)	91.2(2.22)	94.1(1.83)
	LINEAR	92.61(3.02)	97.74(0.67)	96.44(0.77)	94.63(1.18)	<b>97.97</b> (0.78)	<b>89.61</b> (3.05)	91.53(2.34)	<b>94.36</b> (1.69)
	EXTRA	92.36(3.0)	97.74(0.66)	96.41(0.77)	94.6(1.18)	97.97(0.78)	89.47(3.11)	91.51(2.3)	94.3(1.68)
	DELTAR	87.07(4.67)	93.57(4.07)	91.0(4.23)	94.87(1.28)	97.66(1.42)	85.97(2.34)	89.11(3.84)	91.32(3.12)
	DELTAS	92.64(4.49)	<b>97.76</b> (0.7)	96.41(1.25)	<b>94.99</b> (0.92)	97.83(0.99)	88.69(2.69)	<b>91.64</b> (2.36)	94.28(1.92)
512	UPSAMPLE	95.3(1.09)	98.0(0.64)	97.63(0.34)	95.57(0.87)	98.03(0.55)	92.0(1.49)	93.26(1.05)	95.68(0.86)
	PERTURB	95.33(1.2)	97.94(0.6)	97.6(0.44)	95.5(0.91)	97.91(0.55)	92.03(1.78)	93.21(0.99)	95.65(0.92)
	CVAE	95.46(1.12)	97.89(0.62)	97.54(0.43)	95.36(1.02)	97.93(0.7)	91.34(2.17)	92.73(1.19)	95.46(1.03)
	LINEAR	95.39(1.1)	<b>98.0</b> (0.64)	97.67(0.36)	95.74(0.89)	<b>98.04</b> (0.5)	<b>92.61</b> (1.47)	93.66(1.13)	<b>95.87</b> (0.87)
	EXTRA	95.36(1.17)	98.0(0.64)	97.66(0.37)	95.74(0.88)	98.04(0.5)	92.29(1.52)	93.63(1.17)	95.82(0.89)
	DELTAR	95.36(1.74)	97.81(0.69)	97.6(0.44)	95.9(0.97)	97.74(1.02)	90.27(3.44)	92.61(2.64)	95.33(1.56)
	DELTAS	<b>95.66</b> (1.18)	97.96(0.59)	<b>97.8</b> (0.45)	<b>95.91</b> (0.88)	97.91(0.74)	92.26(2.57)	<b>93.66</b> (0.86)	<b>95.88</b> (1.04)

Table 1: IC accuracy on SNIPS dataset in the FSI setting ( $k = 10$ ), reported as *mean* (*SD*).

#	Method	Directions	Distance	Arrival	Duration	Traffic	Event	Overall Mean
No Augmentation		89.61(0.1)	89.94(0.09)	90.56(0.12)	81.74(0.13)	68.5(0.13)	67.39(0.11)	81.29(0.11)
100	UPSAMPLE	89.89(0.27)	93.64(0.87)	92.95(0.57)	84.28(3.45)	68.99(0.49)	<b>76.26</b> (5.41)	<b>84.34</b> (1.84)
	PERTURB	89.82(0.24)	93.58(0.84)	92.81(0.55)	<b>84.81</b> (3.77)	<b>69.15</b> (0.68)	74.07(5.6)	84.04(1.95)
	CVAE	89.91(0.32)	93.46(0.77)	92.7(0.67)	84.45(3.52)	69.11(0.9)	74.94(5.49)	84.1(1.94)
	LINEAR	<b>89.93</b> (0.24)	<b>93.65</b> (0.88)	<b>92.98</b> (0.57)	84.2(3.44)	68.96(0.51)	76.12(5.77)	84.31(1.9)
	EXTRA	89.88(0.27)	93.61(0.89)	92.96(0.59)	84.21(3.43)	68.94(0.46)	75.18(5.34)	84.13(1.83)
	DELTAR	89.64(0.11)	92.57(1.3)	90.79(0.37)	81.72(0.12)	68.48(0.08)	68.63(2.59)	81.97(0.76)
	DELTAS	89.88(0.34)	93.68(0.72)	92.6(0.76)	83.88(3.2)	68.93(0.67)	72.05(5.83)	83.5(1.92)
512	UPSAMPLE	91.93(0.48)	94.58(0.34)	93.99(0.31)	92.56(0.72)	75.84(2.19)	85.27(1.87)	89.03(0.99)
	PERTURB	91.78(0.49)	94.58(0.43)	94.02(0.25)	92.53(0.87)	<b>76.0</b> (2.27)	85.22(1.61)	89.02(0.99)
	CVAE	91.85(0.52)	94.57(0.39)	94.0(0.34)	92.45(0.92)	74.91(2.73)	84.5(1.61)	88.71(1.09)
	LINEAR	92.14(0.66)	94.6(0.35)	94.05(0.32)	<b>92.78</b> (0.67)	<b>76.0</b> (2.49)	86.22(1.7)	<b>89.3</b> (1.03)
	EXTRA	<b>92.11</b> (0.57)	94.61(0.35)	94.04(0.29)	92.72(0.7)	75.79(2.45)	85.98(1.58)	89.21(0.99)
	DELTAR	90.43(0.55)	94.54(0.35)	93.8(0.3)	86.64(4.38)	71.68(1.46)	<b>86.55</b> (1.75)	87.28(1.46)
	DELTAS	91.83(0.47)	<b>94.66</b> (0.4)	<b>94.08</b> (0.24)	92.31(1.45)	75.81(2.1)	86.23(2.08)	89.15(1.12)

Table 2: IC accuracy on FBDialog dataset in the FSI setting ( $k = 10$ ), reported as *mean* (*SD*).

#	Method	Playlist	Restaurant	Weather	Music	Book	Work	Event	Overall Mean
No Augmentation		96.0(1.69)	95.39(1.59)	96.41(1.18)	93.1(1.38)	97.79(0.77)	88.46(1.14)	93.49(0.87)	94.38(1.23)
100	UPSAMPLE	96.0(1.57)	95.87(1.26)	<b>96.51</b> (1.04)	93.19(1.25)	97.83(0.7)	88.63(1.21)	93.7(0.83)	94.53(1.12)
	PERTURB	96.1(1.64)	95.7(1.23)	96.43(1.28)	93.33(1.1)	97.8(0.77)	88.56(1.32)	93.7(0.9)	94.52(1.18)
	CVAE	96.07(1.46)	<b>95.91</b> (1.43)	96.43(1.31)	93.2(1.15)	97.83(0.78)	88.63(1.28)	93.66(0.86)	94.53(1.18)
	LINEAR	96.0(1.57)	95.89(1.26)	<b>96.51</b> (1.04)	93.19(1.25)	97.83(0.7)	88.63(1.21)	93.7(0.83)	94.53(1.12)
	EXTRA	96.0(1.57)	95.84(1.3)	<b>96.51</b> (1.04)	93.19(1.25)	97.83(0.7)	88.63(1.21)	93.7(0.83)	94.53(1.13)
	DELTAR	96.09(1.51)	95.74(1.46)	96.44(1.29)	<b>93.56</b> (0.95)	<b>97.86</b> (0.75)	<b>88.79</b> (1.25)	<b>93.86</b> (0.93)	<b>94.62</b> (1.16)
	DELTAS	<b>96.11</b> (1.52)	95.69(1.44)	96.46(1.29)	93.44(0.93)	<b>97.86</b> (0.75)	88.64(1.18)	93.76(0.89)	94.57(1.14)
512	UPSAMPLE	96.07(1.54)	96.09(1.2)	96.6(1.06)	93.5(1.14)	<b>97.87</b> (0.69)	88.73(1.23)	93.8(0.92)	94.67(1.11)
	PERTURB	96.23(1.6)	96.17(1.23)	96.63(1.13)	93.49(1.03)	97.84(0.72)	88.6(1.3)	93.79(0.98)	94.68(1.14)
	CVAE	96.14(1.46)	96.24(1.18)	96.63(1.06)	93.6(1.08)	97.87(0.75)	88.76(1.29)	93.87(0.98)	94.73(1.11)
	LINEAR	96.07(1.54)	96.11(1.21)	96.6(1.06)	93.49(1.13)	<b>97.87</b> (0.69)	88.76(1.25)	93.8(0.92)	94.67(1.11)
	EXTRA	96.07(1.54)	96.13(1.18)	96.6(1.06)	93.5(1.14)	<b>97.87</b> (0.69)	88.73(1.25)	93.8(0.92)	94.67(1.11)
	DELTAR	<b>96.29</b> (1.52)	<b>96.29</b> (1.34)	<b>96.71</b> (1.1)	<b>93.87</b> (1.04)	97.86(0.75)	<b>89.11</b> (1.22)	<b>94.03</b> (0.89)	<b>94.88</b> (1.12)
	DELTAS	96.19(1.61)	96.2(1.23)	96.69(1.07)	93.61(0.96)	97.86(0.75)	88.84(1.28)	93.83(0.94)	94.74(1.12)

Table 3: IC accuracy on SNIPS dataset in the FSI setting, reported as *mean* (*SD*). The 10 seed examples are upsampled to 100 to train the feature extractor.